DOCUMENT RESUME

ED 323 284 UD 027 640

AUTHOR	De Lone, Richard H.
TITLE	Replication: A Strategy To Improve the Delivery of Education and Job Training Programs.
INSTITUTION	Public/Private Ventures, Philadelphia, PA.
PUB DATE	90
NOTE	69p.
AVAILABLE FROM	Public/Private Ventures, 399 Market Street, Philadelphia, PA 19106-2178.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Case Studies; *Demonstration Programs; Elementary Secondary Education; *Job Training; Models; *Program Effectiveness; *Program Implementation
IDENTIFIERS	*Replication

ABSTRACT

This study draws on literature reviews, interviews, and case studies to examine the viability of replication strategies in education and employment training. Replication is a dynamic process for improving the quality of local services by developing credible knowledge about effective program models, and reproducing those models in a variety of local settings by a process that includes adaptation and evolution while maintaining the results of the original. "Planned replication" by a third-party planning agency appears to be more effective than either legislated replication or locally initiated replication. The effectiveness of the following examples of planned replication of education and job training programs is reviewed: (1) the Comprehensive Competencies Program (CCP); (2) the Summer Training and Education Program (STEP); (3) Jobs for America's Graduates (JAG); (4) Experience Based Career Education (EBCE); and (5) the Coalition of Essential Schools. The following conditions are necessary for successful replication: (1) evidence of the effectiveness of the model to be replicated; (2) reasonable input costs; (3) availability of a qualified replicating agent; (4) an identifiable market need for the program or service to be replicated; (5) choice of a replication strategy appropriate to both the product and the institutional context in which the replication will occur; and (6) provision for continuing evaluation. A list of 84 notes and references are appended. (FMW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Public/Private Ventures

P/PV

ED323284

Replication

A Strategy to Improve the Delivery of Education and Job Training Programs

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

---

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
N. Jaffe
Public/Private Venture
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UD 027 640

Summer 1990

BEST COPY AVAILABLE

Public/Private Ventures is a national, not-for-profit corporation that designs, manages, and evaluates social policy initiatives aimed at helping people whose lack of preparation for the work force hampers their chances for productive lives. P/PV's work is supported by funds from both the public and private sectors.

Rex D. Adams
Chairman
of the Board

Michael A. Bailin
President

Gary Walker
Executive
Vice President

Board of Directors

Rex D. Adams
Michael A. Bailin
Judy Barker
Angela Glover Blackwell
Roy J. Bostock
Alan K. Campbell
Hillary Rodham Clinton
Alonzo A. Crim
Linda Darling-Hammond
Alice F. Emerson
Sandra Feldman
Harold Howe II
Siobhan Oppenheimer-Nicolau
Marion Pines
John W. Porter
Mitchell Sviridoff

Corporate Associates

American Express
Philanthropic Program
Ameritech Foundation
ARA Services, Inc.
Borden Foundation
BP America
CIGNA Foundation
Colgate-Palmolive Co.
Dean Witter Reynolds, Inc.
Exxon Corporation
General Motors Corporation
Independence Blue Cross
Mobil Oil Foundation
Penn Mutual Charitable Trust
Pfizer, Inc.
The Prudential Insurance Co.
of America
Scott Paper Co.
Tenneco, Inc.
Union Carbide Corporation
United Technologies

Research Advisory Committee

Frank Furstenberg
Robinson Hollister
Henry Levin
Frank Levy
Anita Summers
Marta Tienda

Public/Private Ventures
399 Market Street
Philadelphia, PA 19106-2178
215/592-9099

P/PV

Replication

A Strategy to Improve the Delivery of Education and Job Training Programs

by Richard H. De Lone

Summer 1990

CONTENTS

Page

I. INTRODUCTION. 1

II. DEFINING REPLICATION. 5

III. PLANNED REPLICATION: ISSUES AND EXAMPLES. 17

IV. LESSONS FROM EXPERIENCE: THEMES AND VARIATIONS OF REPLICATION. 37

V. CONCLUSIONS: WHEN TO REPLICATE AND WHY. 53

NOTES. 59

REFERENCES 61

FOREWORD

The most persuasive reason for doing research and demonstrations in the social policy arena is to apply widely their findings about what works. Over time, this process should make social programming more efficient, and should reduce social problems. Unfortunately, this scenario remains largely theoretical.

The art of mounting demonstrations and doing credible research on social programs has become more sophisticated over the past two decades. We have seen notable advances in the science of "finding what works." But we have not seen equal growth in our ability to put these lessons into broader practice. There have been only a few successful efforts to install the specific practices of successful programs around the country.

In the last three years, P/PV has undertaken--with Department of Labor and philanthropic support--a replication of its successful summer program STEP (Summer Training and Education Program). Over that period, STEP has grown from its five demonstration sites, enrolling slightly over 500 youth a year, to 73 sites and more than 7,000 youth. The new sites have performed as well as the demonstration sites, so we are confident the replication is a true one. To accomplish that, we found replication to be time- and resource-intensive, involving repeated rounds of training, on-site assistance, monitoring and assessment, and far from cheap--at least by the standards of social program technical assistance. Replicating STEP costs from $15,000 to $20,000 per site.

We also found little precedent to guide this effort and no alternative to doing the training for replication ourselves. Yet, as a demonstration/research intermediary, we know that our organization cannot undertake too much replication work without diluting our basic mission or our reputation for credible research.

Ensuing discussions with other professionals in the social policy field have disclosed widespread concern about the replication issue. A few felt it to be so fundamental that its resolution should take precedence over any further demonstration and research work. Nevertheless, we uncovered little systematic thinking about replication. We decided, therefore, to undertake a multistage investigation, with the first stage being an exploratory paper that would lay out some of the critical concerns, operational and structural, that seem to underlie the lack of replication experience and thinking.

This report by Richard de Lone is the product of that first-stage exploratory effort. As de Lone's report indicates, there are examples of successful replication. But most of them had no prior evidence of program effectiveness--in a few cases that evidence came later, after replication (Job Corps, Head Start). Most show such a wide variety of practices among the different jurisdictions that the original model is scarcely recognizable. And few are able to document their ongoing success. In short, the evolution from theory to evidence to widespread practice has not worked very well.

The lack of a strong connection between evidence-gathering and field practice would be a niggling detail if our major social initiatives were working effectively. But they are not. In fact, a few critics argue that some social initiatives prolong dependency and worsen social problems; more analysts maintain that these initiatives are usually of little consequence. The credible evidence we do have often supports the latter claim. No one contends that either our welfare or employment/training initiatives, to take but two examples, are significantly reducing major social problems--and the reason has as much to do with program effectiveness as it does with lack of funding.

Thus, it is critical that we make more focused efforts to ensure that what we learn gets put into widespread practice, and that we understand how that is done. The atmosphere of institutional change that currently prevails in the social policy world makes our understanding of the practice of replication particularly vital, since, to be effective, large institutional change initiatives must ultimately be composed of specific program components.

An obvious question is why replication has not received more concerted attention--especially in a country whose commercial sector has refined to a science the practice of replication and franchising. As de Lone's paper points out, the very different status of replication in the social policy and commercial worlds has much to do with the very different funding structures and financial incentives that prevail in those worlds. In the commercial world, a successful initiative is copied widely and instantly. Economic gain is the personal incentive, and lending institutions prefer to go with a winner.

Philanthropic funders in the social policy world, however, tend to like new, creative initiatives, not repetitions. Public funding sources are typically not generous with the staff training and consultant monies required to bring on effective practices from other jurisdictions. State and local administrators have no clear bottom line for most social programs; they are rewarded, rather, for their inventiveness. With limited demand and financing, and few rewards, the unsurprising result is that few agencies or individuals even try to replicate their programs elsewhere, or replicate the programs of others in their own jurisdictions.

Certainly, P/PV does not believe that replication alone will solve the social problems facing us. Social policy issues are substantially more complex than hamburgers, and we do not foresee a chain of McDonald's-like social programs dominating the landscape.

However, we do think that replication is a vastly underused tool for creating better programming. Common sense and local creativity will simply not be enough to mount multicomponent programs with specialized curricula in education, life skills and skills training. Replication's usefulness is more important now than ever, as social programs grow more complex to meet more complex social problems.

Gary Walker
Executive Vice President
Public/Private Ventures

I. INTRODUCTION

The idea seems simple, appealing and rich with possibility: find out what works in education, job training or other social and human services and replicate it everywhere the same need exists. The appeal is undeniable, and this paper will argue that the idea has considerable, though largely unfulfilled, potential. But simple it most certainly is not.

There are three key concepts here: "what works," "replication" and "same need," each of which is complex and problematic. Each poses difficulties in definition, measurement and implementation; each inhabits the realm of applied social science, an eminently political realm where diverse and often conflicting values, interests, capacities and structures characterize and complicate situations; and, finally, each unfolds over time in environments that are fluid and evolving. Replication, in short, is a strategy for change that itself must adapt to diverse and changing circumstances.

This paper is an effort to both amplify and clarify the issues involved in replication. Its basic question is whether the appeal and potential of the idea can be realized without foundering on complexities that have, in fact, often sunk simplistic efforts. The question is important to ask for three basic reasons:

1. There is little point in investigating "best practice" through research, pilot programs, demonstration programs or program evaluations if one is unable to replicate, adapt or transfer lessons about effective programming from single to multiple sites.

2. Improvement of local implementation has become widely recognized as a critical issue (some would argue the critical issue) for improving returns on social investments, and replication is one way to attempt to improve local practice.

3. The replication process itself, while often envisioned by funders and policy planners as an eventual goal, has not been the topic of much systematic investigation or attention.[1] The process is taken for granted or assumed to be easy once the information is available--in part explaining why systematic replication is urged more often than it is achieved.

Interest in replicating effective programs springs from many sources, as indicated by a wide variety of publications that attempt to describe "what works"--with the implicit or explicit hope that what works will be replicated. Less widespread, but still quite varied, are systematic efforts to execute replications. Think tanks, foundations, governments at all levels, the military, school/business partnerships and others are among these sources of interest and activity. Replication is particularly appealing to those who believe we know more than we use; examples are Schorr (1988) and Walberg (1989). Walberg argues that more than 7,000 educational interventions worldwide have been shown to produce results, but only a handful are in evidence in most schools.

Perhaps the most compelling a priori case for replication emanates from a concern for improving quality in programs heavily funded by state or federal governments but implemented by relatively autonomous local units--such as education (state-funded), or employment and training of the disadvantaged (federally funded). Indeed, much of the interest in applying social science to social problem-solving--of which replication is one approach--grew out of concern that the huge increases in domestic program spending during the Great Society era were not effective (Rivlin, 1975).

Replication also has detractors and reasoned skeptics. There are those who--unlike the Walbergs--argue that we know so little about what actually works that there is little cause to replicate. While there is plenty we do not know, the contention of ignorance is palpably false unless one adheres to irrationally strict standards of proof. There are others who resist--indeed, resent--the idea of replication. These include practitioners who have suffered from heavy-handed, inept, top-down implementation of programs, as well as victims of the "not invented here" syndrome. Urban school teachers, for example, many of whom feel battered by central mandates, are among those for whom replication is "the dreaded R word."[2] But replication need not mean forced compliance. Nor need it be the foe of individual creativity.

Replication is not, of course, the only way to improve the implementation of social programs. Adequate funding, well-conceived legislation and regulations, performance incentives, sound planning, good evaluation, and a wide range of training and technical assistance to strengthen management or staff capabilities are among other ways to strengthen implementation. These are certainly as important and more broadly applicable than replication. In public education, many argue that basic structural reforms designed to enhance teacher professionalism, or increase parental choice, or transform the nature of curriculum and pedagogy, are what is most needed. These strategies may or may not be requisites for effective implementation, but they need neither negate nor compete with replication strategies. Rather, replication can complement other strategies, for whenever the question of "best practices" arises, as it almost inevitably does (or should), there is at least a potential for replication to complement other strategies for implementation.

This paper draws on literature, interviews and case studies to examine the viability of replication strategies in two fields, education and employment training. While both have similar tasks--the development of human abilities--they represent interesting contrasts in institutional context.

To preview conclusions, I will argue that replication is in fact a strategy with considerable merit; that carefully planned replications, duly adapted to local circumstance, have met with considerable success in a few instances; and that the requisites of effective replication are well enough understood to warrant placing much more emphasis on this aspect of improving social programs. In particular, I will argue that replication is most likely to succeed if pursued in an entrepreneurial, market- driven style by an independent replicating agent (as opposed to by top-down or bureaucratic directives). Replications will not work everywhere, but they can fill often large niches in the public policy marketplace. The task of governmental agencies and funders is to determine where those niches exist and to invest in the capacity of organizations to be effective replicating agents.

At the same time, however, caveats abound. Replication is no panacea; it is a strategy for modest incremental improvements, not great leaps forward. Replication is easiest to do if one is augmenting services and harder/less proven as a strategy for changing existing institutions or professional practices; different circumstances and purposes call for different approaches, and any approach will be a multistep process in which failure at any step can undermine the whole endeavor. There may also be problems involved in producing valid evidence to determine that the replicas work as well as the original. The replication process requires long timelines in a world dominated by the short attention spans of political creatures. Replication is not immune to the vagaries of politics, professional competence, institutional turbulence or resource inadequacies--but neither is any other strategy for improving implementation. Further, one problem to which planned replication responds--the highly decentralized, fragmented and unstable world of public programs--is also a major obstacle to replication.

Caveats aside, however, planned replication, skillfully executed, has some distinct advantages for addressing a range of critical issues in implementation. These include taking the opportunity to develop and refine program design through successive iterations; transferring and spreading knowledge of sound programming into practice; building bridges between research and practice; strengthening the quality of programs whose success depends on decisions, interactions and configurations of resources beyond the control of any individual practitioner; and assuring continuity of effort and focus, whose absence is often the biggest stumbling block to improving implementation in the hurly-burly world of public policy. Finally, in a realm too often lacking substantial focus on a "bottom line" of results, replication offers a way to incorporate both implementational rigor and performance-forged output standards into social programming.

To reach these conclusions this paper will:

1. Define replication and discuss the pros and cons of different modes through which it may occur: legislated, locally initiated and planned replications.

2. Present some examples of planned replications in education and job training programs in the last 20 years.

3. Discuss issues raised by the cases, including:

    - Research criteria and uses
    - Critical success factors in planned replication
    - Strategic options for planned replication
    - Technical issues

4. Discuss the role of government in promoting replication.

5. Summarize findings and address the questions of when and why to replicate.

II. DEFINING REPLICATION

A first approximation of a definition of replication is "*a process by which programs of credible effectiveness are reproduced and embodied widely in practice*."[3] This definition views replication as a two-part process, consisting of 1) the development of credible knowledge about effectiveness and 2) the reproduction of programs that embody that knowledge. We will address the standard of credible knowledge later, but the meaning of "reproduced" deserves immediate attention.

Clearly, cloning, slavish imitation or other forms of "exact" reproduction are neither feasible nor desirable for most social programs. But a "replica," as the *American Heritage Dictionary* tells us, need only be a "close reproduction." The key question is how close--i.e., with what degree of fidelity must the model be copied? The best answer is: closely enough to replicate (achieve or surpass) the results of the original.

This *ex post facto* reply is important. As Walter Williams put it some time ago, "The objective is performance, not conformance." (Williams, 1976) However, the *ex post facto* reply is of little immediate use to the replicator, who still needs to know how closely to reproduce the original to get results. There is a general answer to the question: it must be close enough to preserve the "essential integrity" of the model. This is not to beg the question, but rather to point out the need for thinking systematically about the answer in each and every instance of replication, for there are few if any invariable rules.

Some models are quite general, consisting primarily of a few guiding principles, and rely on practitioners to exercise considerable judgment to activate those principles. Others are tightly specified. Some models are specific only about organizational structure; others specify program activities and techniques. Some models may require "high-fidelity" implementation in early stages, but permit, encourage and, indeed, become improved over time by evolution, variation or experimentation.

Occasionally, research may yield definitive evidence that certain aspects of a model are crucial to the desired outcome. More often, the research will at best give hints about which elements in the "black box" of the program are most critical to positive results. Such research limitations blur the imperative for high fidelity replication and should give the honest replicator pause for modesty.

Some models, in their "ideal" form, simply will not be feasible to implement in a particular context without a fair amount of adaptation. In other instances, adaptation may be warranted as a trade-off to obtain a sense of local ownership and commitment.

How much adaptation? The answer depends on a thoughtful assessment of what is needed to preserve "essential integrity"--and results. How much evolution? The answer is the same. The judgement-prone nature of replication points to the need for a skilled agent of replication who can make those judgments (another topic to which we shall return).

For these (and other) reasons, it is best to define replication as a dynamic process that, in different instances, places and times, will include mixes of fidelity, adaptation and evolution of the original model. The right mix will depend on systematic and informed judgments about what is essential to the integrity of the model, as well as confirmation from ongoing evaluation and documentation of results.

ALTERNATIVE MODES OF REPLICATION

Given this second approximation of a definition, replication may be attempted in very different modes. Three primary modes are legislated, planned and locally initiated. We will examine legislated and locally initiated modes briefly, before the paper turns to its main topic, planned replications.

Legislated Replication

If replication is to be developed as a strategy for improving local program implementation, legislation and state and federal governments will, perforce, have an important role to play in stimulating it, as we will discuss. But it is unlikely that they will play that role by legislating the adoption of particular program approaches. There are some examples of efforts to develop and incorporate "credible knowledge" into legislation that is fairly prescriptive as to program design and implementation, but it is *a priori* clear that the legislative process is only occasionally amenable to actually specifying program models. Most legislation results from a series of compromises among competing interests and visions, and research findings are of marginal importance, at best, to the politics of social legislation. For the most part, state and federal governments lack both the mandate and the staff capacity to engage in the kinds of training, technical assistance or site monitoring and evaluation which, as we shall see, successful replication appears to require.

Further, the political process is impatient, and the time horizons of legislators are usually fairly short--often extending no further than the next election. By contrast, the patience and time frames required to develop and test a program model, evaluate its effectiveness, and get it up and running well in a range of jurisdictions are usually long. In our federalist system of governance, major legislation, for reasons both substantive and political, usually requires fairly broad guidelines permitting considerable variation in program design and process from jurisdiction to jurisdiction. Specificity is focused on targeted population, allowable activities, cost parameters, allocation of dollars and, sometimes, performance outcomes--as in such pieces of major legislation as Chapter 1 of the Education Consolidation and Improvement Act, or the Job Training Partnership Act. How best to design and implement programs supported by such legislation is largely a matter of local discretion.

On occasions when major legislation attempts to be fairly prescriptive as to programming requirements, the results are often discouraging. An example is the spate of "educational excellence reforms" passed by state after state in the early 1980s in the aftermath of A Nation at Risk. Legislators, we can assume, felt they had "credible knowledge" that specifying sequences of courses, and/or promotion standards, and/or requirements for graduation would improve schooling. In essence, most of the excellence reforms were crude models for high school programs. Recent research suggests that while these "reforms" generated headlines and caused a blizzard of paperwork and more than a few headaches for local school administrators, they produced very little change. For the most part, school officials changed the names of their courses to comply with the law, and continued doing whatever they had been doing before. Even if one believes that the excellence reforms were well-conceived--a dubious proposition--they lacked any mechanism to produce the desired changes (Clune, 1988).

There are, nevertheless, some examples of federally legislated efforts to engage directly in replication, more or less as we have defined it--i.e., involving both the development of credible knowledge and the adaptation of proven programs to multiple sites. We will briefly review four efforts which are instructive both for the lessons they convey about the replication process and for the lessons they contain about the constraints on legislated replication. These efforts are the U.S. Department of Education's National Diffusion Network (NDN); the 1977 Youth Employment and Development Projects Act (YEDPA), administered by the U.S. Department of Labor (DOL); the Job Corps, also administered by DOL; and the Head Start program.

National Diffusion Network (NDN)

The NDN, initiated in the early 1970s as part of the now-defunct National Institute of Education, is a small but interesting approach to developing and replicating effective educational programs. Educators submit research findings on programs to the NDN, which assembles expert panels to review the research. If the panel validates the research, the program is certified for dissemination. Small grants are available from NDN to replicate the program in other settings, frequently with technical assistance provided by the person or persons who initiated the original model. Unfortunately, systematic evaluation of the effectiveness of this replication strategy is not performed. Inclusion of some NDN projects in a large study of educational innovation suggests that results are mixed, with the successful transfer of knowledge depending considerably--as one might expect--on such variables as the extent of local administrative support and the quantity and quality of technical assistance offered to replicators (Huberman and Miles, 1984).

On the face of it, the NDN suggests a creative way for federal or state entities to honor the political realities of pluralism, by looking to local initiative to create models, while also disseminating effective strategies across a range of jurisdictions. NDN has too small a budget (circa $14 million a year) to develop and test fully the approach. More rigor goes into validating research than into dissemination strategies. While Congress has continued to support the program on a small scale, its expansion, according to a senior department official, has always been limited by opposition from locally rooted educational groups, which lobby to get federal dollars directly. Thus, while the NDN has considerable intuitive appeal as a federal (or state) strategy for replication in a decentralized service delivery system, the politics of that system curb its potential.

Youth Employment and Development Projects Act (YEDPA)

YEDPA, passed early in the Carter administration, was a highly unusual piece of legislation in that the Congress set forth in it a fairly detailed agenda for research on "what works," indicative of both the uncertainty and the frustration many lawmakers felt about how to redress the seemingly intractable problems of growing youth unemployment, especially in inner cities. While the lion's share of YEDPA funds were distributed to localities under formula, a substantial amount remained under the control of the Department of Labor for research and experimentation--the "knowledge development" program.

Perhaps because the federal government is the major source of funding for job training programs for the disadvantaged, there has been a fairly strong tradition of research and evaluation in DOL's Employment and Training Administration (ETA). Especially in the late 1960s and throughout the 1970s, this included support for large-scale demonstrations which ETA hoped would, among other things, produce knowledge about effective programming which could be incorporated into legislation--a replication strategy writ large. Howard Rosen, the retired head of ETA's research division, who oversaw many of its most ambitious research efforts, believes that research has influenced legislation--albeit in ways that are often "hard to trace," since research is at best one input into the legislative process. But for the most part (we will address one major exception later), that influence has taken the form, again, of broad principles and guidelines, not either highly or loosely specified program models that are the stuff of replication.

The history of YEDPA, whose congressionally mandated knowledge development agenda placed a heavy emphasis on testing program models and in one case actually called for a replication experiment, suggests some of the reasons why. First and foremost, the timetable for knowledge development was simply too short. YEDPA had two years to produce results. A little reflection suggests that it is amazing any results were produced at all. For it takes a minimum of three to six months to get the money flowing after Congress passes a new bill; it takes another three to six months (moving at a rather frenzied pace not highly compatible with the careful thought that should go into research programs) to get programs in place to evaluate. Ideally, programs should be running for a while--perhaps a year--to pass through learning curves and get the bugs out. (There is no point in heavy evaluation investments in a program that isn't running smoothly.) Impact evaluation, then, might begin with the second year of program implementation. In both education and job training programs, meaningful assessment of impact on participants almost always requires postprogram follow-up studies of a year at the very least. Add these up, and one has a four- or five-year timetable as a reasonable minimum, compared to YEDPA's two-year timetable.

Second, in the pressure to develop experimental programs and produce research results (in a truncated time frame), little if any systematic attention was given to the strategy for replicating programs if the research demonstrated their worth. And, third, even if such attention had been given, it is dubious that it would have been implemented. YEDPA was part of the Comprehensive Employment and Training Act (CETA). By the time the results from YEDPA began to come in--even with a hurry-up timetable--the Carter administration and CETA were on their way out, to be replaced by a new job training law (JTPA) and a new administration. Provisions of the new law made it next to impossible to implement many of the program models tested under YEDPA, and youth employment programs were hardly high on the priority list of the new administration.

Not all the YEDPA research was in vain. Many practitioners involved in it have continued to build subsequent work at least in part on the base of knowledge and experience developed then. But there was no orderly transfer of research findings into national legislation.

Job Corps

One piece of legislation that has survived as a program over several administrations is the Job Corps, which provides education and training for disadvantaged, hard-to-employ youth in residential settings. Initiated in the Great Society years, the Job Corps is one of few directly funded, federally managed social service programs (outside of the Department of Defense). While not planned as a replication, per se, it at least partially fits our definition of one. While there is some variation from site to site, the basic Job Corps model is pretty standard. Most sites are operated by private organizations under contract to DOL, and one contractor operates multiple sites; thus, a strong strategy for very precise replication is in place.

The Job Corps was not initially developed through an R & D process, but it has been frequently evaluated and was the subject of an extremely rigorous empirical study funded by ETA in the 1970s. This study produced strong evidence that, despite its high participant cost, the benefits of the program in terms of increased employment and earnings of its participants, and decreased social costs make it cost-effective (Mallar, C. et al., 1982). It is generally believed that this strong research evidence--combined with astute cultivation of politicians by Job Corps officials and advocates--saved the program from Reagan administration cuts. The Job Corps, then, suggests the viability of legislated replication, but it is also a very special case that tends to be the exception to the rule. Direct federal operation of programs is uncommon and unlikely to increase. There is no influential state or local interest group claiming it can run Job Corps better than the "feds." Since the concern of this paper is the viability of replication in a decentralized system, there is little to be learned from the Job Corps experience--except that the combination of stable funding, strong research, and competitive use of contractors to deliver a model can produce quality programming.

Head Start

A further illustration of the possibilities--and limitations--of "legislated replication" is Head Start. At first glance, Head Start resembles a successfully legislated replication. Not only must Head Start programs meet typical federal requirements for target population (income and age) and cost, there are also guidelines about services to be provided (e.g., education, nutrition and health care), parental involvement, and staffing (use of indigenous paraprofessionals) that begin to describe something resembling a program model. While there is considerable diversity in Head Start, one can visit a Head Start center in Boise or Brooklyn and rapidly recognize it as a Head Start program. Further, it is widely believed that the effectiveness of Head Start has been proven by research. In reality, as we shall see, all that research has demonstrated with any conclusiveness is that some early childhood programs have had some lasting effects on children. There is no conclusive evidence that the Head Start legislation has produced such effects, and no certainty as to whether some models for Head Start (e.g., different curricular approaches) are more effective than others. In fact, as discussed in a few pages, efforts to develop knowledge about what makes Head Start work constitute a cautionary tale about the difficulties of developing credible knowledge. Yet, because a few bits of evidence suggest Head Start could be so valuable an intervention, it remains a prime case in the argument for greater effort at discovering "what works" and replicating it.

Conclusions

Without support from legislation and the administrative arms of state and local government, it seems unlikely that replication can be developed as a major strategy for improving local implementation. However, as the nature of the legislative process and the examples we have cursorily reviewed suggest, achieving that support on a large scale is more complicated than simply discovering what works and writing program parameters into new laws. Well-crafted legislation is important to the possibility of replication, but--as will be discussed in the final section of this paper--in all but rare instances, replications cannot be directly legislated. Instead, legislation is more likely to be effective in stimulating replication through incentives and investments in the capacity to replicate.

Locally Initiated Replications

An argument frequently advanced against replication is that every community (school, agency) is different--from place to place and from year to year. Once replication is defined to include adaptation, this argument loses considerable force. But the contrary is also true: replication can be and often is a strategy initiated by local entities to mount a new service or modify an existing one. Casual observation suggests that such local replications are widespread, though I know of no systematic study of the process.

Indeed, the motives for replication are fairly obvious and varied: people are not satisfied with the current state of affairs and look for a better way to do things; someone with little experience is told to start a program quickly and looks for a template; an administrator goes to a conference and hears about a program that strikes his or her fancy. For these and other reasons, replication happens every day. Sometimes, the copy may be quite exact; at other times, quite loose. Sometimes, the process is thoughtful and sometimes it is a "quick fix." Sometimes it works, and sometimes it does not. There is no reason to think that the rules governing locally initiated replication are much different from the rules governing program planning and management in general--or that the outcomes of locally initiated replications are much different from those of local programs in general.

This does not mean, incidentally, that effective local replication cannot be promoted. It can be done by improving the capacity of local practitioners to use research and to adapt successful approaches to their situation. In my own work with schools engaged in attempting to restructure and provide teachers the time, authority and resources to get involved in program planning, I have seen numerous examples that illustrate the point. Enabling practitioners to use research is highly desirable, in my view, but it does not quite address the problem of taking what is known to be an effective program model and replicating it.

If legislated replication is both a weak form of replication and hard to do, then local replication may be a relatively strong form of replication that is easy to do--but not necessarily easy to do well. It is inherently difficult to control the frequency or quality of local replications.

Accordingly, if replication as we have defined it holds promise as a means to improve the quality of local programming, it will require a third approach: planned replication, in the hands of some mediating agency positioned between the Potomac and the local service agency with the staying power to outlast changing political fads; the skills to operate across the diverse and slippery terrains of policy, research and programming implementation; the small "p" political skill to work across various jurisdictions; and the sense and sensitivity to maintain the essential integrity of a model while adapting to local circumstances.

PLANNED REPLICATION: THE LESSONS OF HEAD START

Before reviewing some examples of planned replication and assessing this strategy, it may be useful to consider a couple of early examples: the Head Start and Follow Through Planned Variations. The former was an effort to develop a planned replication strategy as a means of improving an existing piece of legislation that, as noted above, can be considered in some ways a successful example of legislated replication (one can tell it's Head Start in Boise and Brooklyn). The latter incorporated phase one of a planned replication--the development of credible knowledge--right into the law.

The Head Start legislation, initially funded under the Office of Economic Opportunity as part of the "War on Poverty," has always required certain services, parent participation and staffing patterns. These broad parameters of a model were based on a combination of research on human development and what the framers of Head Start saw as critical components of good existing early childhood programs. However, programs could comply with the guidelines and still offer a wide variety of educational approaches and developmental activities, ranging, for example, from highly structured curricula designed to teach reading, to

Piagetian models attempting to stimulate general cognitive development, to highly child-centered approaches, etc.

Early evaluations of Head Start often (but not always) found positive initial results in cognitive as well as social and personal growth. But most of those same evaluations also found that, by the third grade, gains experienced by participants in early childhood programs had disappeared (Datta, 1977). These findings led to a federally supported effort to investigate which Head Start models were most effective at producing various outcomes, and whether some produced longer-term benefits. The Head Start Planned Variation experiment, designed in 1969 and conducted over the next five years, attempted to test 11 different "models," from highly specified curricular models to an open-ended "Enabler" model in which the "treatment" was to let each community design its own program.

This effort was paralleled by a related experiment, the Follow Through Planned Variations.[4] Follow Through was designed to test models of primary school programs to see which would best sustain Head Start gains. Together, these large-scale efforts arguably constituted the high water mark of efforts to apply social science to the development and dissemination of effective public education programs. But the waters did not rise very far. Neither yielded definitive results, for reasons that primarily had to do with difficulties in designing and researching "models" and, secondarily, difficulties in implementation.

As Rivlin and Timpane (1975) summarized the lessons, the models were developed too rapidly and were often poorly specified and/or only partially implemented. The goals and objectives of the models were often vague and lacked in specified criteria for effectiveness. Indeed, in the case of Follow Through, the idea of planned variation itself was a Trojan Horse, with service delivery the real interest of the program's framers (Elmore, 1975). Nor was there "a firm idea of just how and when the hoped-for effects were to appear" (Rivlin and Timpane, 1975b: 13).

Lacking clear criteria and faced with time pressures to produce data, existing tests were used as output measures, though these tests were not always appropriate to the objectives of the model and were "singularly ill-suited for use as discriminators among curricula" (ibid.: 14). Further, neither Planned Variation experiment was truly experimental: they did not use random assignment either of pupils to control and treatment groups, nor of models to sites. The best research design would have done both (Datta, 1975), although random assignment of sites is a virtual impossibility (Elmore, 1975). The costs--in terms of resistance to the model--would probably cancel out benefits in terms of research purity.

Equally important, neither the Head Start nor the Follow Through Planned Variations were conducted with either clear objectives or a strategy for assuring faithful replication. In fact, "lack of knowledge about what actually went into the educational process in the various sites and classrooms makes it difficult to draw conclusions about the outcomes" (Rivlin and Timpane, 1975:14).

With benefit of 15 years of hindsight, there are a number of lessons to be derived from these planned variation efforts. In one sense, they can be dismissed as fairly naive attempts at applying social science techniques to a process of developing, testing and disseminating effective program strategies. They underscore the complexities of applied research, of model development, and of implementation. Call them "learning experiences"--as indeed they were. In another sense, they illustrate the difficulties of applying a rational approach (which replication is) to improving quality in legislated programs. As an effort to move beyond regulations (compliance-oriented replications) they were at once commendable, among the largest efforts ever, and also quite disappointing. Indeed, their failure to produce definitive information was one factor (though not the only one) which reinforced skepticism about the strategy and led most subsequent federal education research away from the development of effective program models into other areas: basic research on learning, and program audits of federally funded activities designed to answer the questions nearest and dearest to a legislator's heart: Did the money get out? Did it get to the people who were supposed to receive it? Did the program look anything like we thought it would? ("Did it work?" finishes a poor fourth.)[5]

Twenty years later, however, the question of Head Start's effectiveness--and more particularly, of whether some approaches are more effective than others--persists. Early childhood mavens and some surprising allies in corporate America (e.g., CED, 1987) are prone to argue that longitudinal research has proven the case for early childhood programs. These claims are most often bolstered by citing the long-term High Scope study of one small carefully designed program (Berruetta-Clement et al., 1984). But it takes a heroic leap from the findings of one study to the recent conclusion, found in an economic report of a large bank, that "every dollar invested in preschool education now returns $4.75 in special education, welfare and prison costs"--a conclusion asserted with great certainty and no footnotes (First Pennsylvania Bank, 1989). In fact, a "meta-analysis" of 210 Head Start studies, recently funded by the Department of Health and Human Services, concluded that while Head Start produces initial gains on a wide range of indicators, all but a few gains, found in only a few studies, wither away by the third grade (McKey et al., 1985).

Schweinhart and Weikart (1986) from the High Scope Foundation, whose longitudinal study of the Perry Preschool Project has yielded the strongest evidence of lasting preschool effects, criticize the meta-analysis on a number of grounds. (Some criticisms appear to have merit; others suggest they did not read McKey et al. carefully).[6] One major criticism they offer is that the meta-analysis failed to distinguish between different types and quality of programs. This is not a trivial issue, since the Perry Pre-School Project itself was not a Head Start program and had both higher per pupil costs and a lower ratio of pupils to professional staff than Head Start generally provides, as well as a particular Piagetian approach to the educational program.

But this simply brings us back full circle: the state of knowledge about the effectiveness of different Head Start models today is little better than it was 20 years ago. The problems encountered by the Head Start (and Follow Through) Planned Variations remind us that there is nothing simple about applying social science to program improvement. But the recurring uncertainty about which Head Start programs "work" suggests the desirability of continuing to try. The examples of planned replication which we will consider next suggest there may be merit in doing so.

III. PLANNED REPLICATION: ISSUES AND EXAMPLES

We have defined replication as a dynamic two-stage process for improving the quality of local services by:

1) Developing credible knowledge about effective program models, and

2) Reproducing those models in a variety of local settings by a process that permits adaptation and evolution while maintaining the "essential integrity" of the model, an integrity whose ultimate measure is results.

Further, we are focusing on planned replications, guided by a replicating agency with the requisite combination of research, implementation and (small "p") political skills as well as the staying power needed to carry replication through its multiple stages.

WHAT IS CREDIBLE EVIDENCE?

As the definition implies, the first step is deciding what constitutes credible evidence. Volumes have been written on this question, which goes to the root of social science's claim to have a role in social problem-solving. At one extreme are the hard-core empiricists who argue that virtually nothing is credible but experimental research, randomly assigning participants to treatment and control groups and testing for statistically significant differences in outcome. Something close to this extreme was represented by the National Academy of Science panel, which discounted the vast majority of research funded by the U.S. Department of Labor on grounds that it was methodologically flawed--i.e., did not use random assignment (Betsey et al., 1985). At the other extreme are those who, like the devil citing scripture, use the incantation "research shows" to give a veneer of certainty to their own beliefs. The latter are meretricious and the former too limited.

There was no control group study of James Comer's celebrated transformation of the King School in New Haven from one of the lowest to one of the highest achieving schools. But the dramatic turnaround and the knowledge that the school's demographics changed little are sufficient to convince anyone that Comer's efforts paid off (Comer, 1980). (It remains to be seen, as Comer attempts to replicate his process in other sites, whether the process--or the involvement of the man himself--was the key to success.)

Similarly, Theodore Sizer's "Coalition of Essential Schools" employs nine key principles that are derived as much from extensive field observation and a systematic effort to apply common sense and experience to the design of schools as they are from any empirical research. One principle, for example, is that no high school teacher will instruct more than 80 pupils. "It does not take a control group study to know that if students are not known in the school, they are not engaged by the school," Sizer argues,[7] explaining the rationale behind this attempt to make teacher-pupil contact in high schools less impersonal.

As Marion Pines, a leading practitioner and authority on job training programs, puts it: "Program operators need the best available information to make decisions, but none of us have the luxury of waiting for perfect information."[8]

Experimental studies of program effectiveness are highly desirable. They produce the greatest degree of confidence in the efficacy of a model and provide a firm peg upon which to hang one's policy hat. They help to reduce uncertainty. But, as a rule, the more narrowly specified the question, the greater the certitude of the answer. Policy-makers and program planners must ask broad questions. In fact, as Lindblom and Cohen argue persuasively in their useful book, Usable Knowledge: Social Science and Social Problem Solving (1979), social science is more than experimental research, and it must be more broadly defined to be useful. Their definition includes a wide range of types of inquiry from "the work of seminal minds like Marx, Freud and Adam Smith" to "statistical manipulation and analysis of social data" to "systematic professional speculative thought," underscoring throughout the qualifier "systematic"--as opposed to "occasional and superficially practiced activities." They stress that useful policy analysis will inevitably be a blend of social science inquiry, "ordinary knowledge"--the domain of common sense and experience without which the policy question cannot even be framed--and "casual analysis."

To make the case for replication, such a combination is imperative. Experimental research may establish that a program is more effective than no program at all--the most common research strategy. But policymakers must also choose between a variety of possible programs. In the real world, it is almost unimaginable that experimental research will determine which is the best alternative. Not only is it difficult to mount planned variations, but it is impossible to be sure, without applying nonstatistical judgments, that all the reasonable alternatives have been tested.

Finally, there are technical reasons why planned variation experiments are unlikely to yield definitive results. Experience shows that, typically, the empirically validated benefits of program models are modest, even when control group members receive no service. In a planned variation experiment, the likelihood of finding statistically significant differences among models (with similar purposes) is less than it is in treatment/no treatment studies. In fact, as the number of variations grows, the likelihood increases that a finding of superior benefits will occur as a matter of chance--what statisticians call a false positive or Type I error. Mosteller (1975) points out that:

> In an experiment including 10 planned variation pilot studies, one of them should stand out as far as 1.5 standard deviations [e.g., the difference between a 500 and a 650 on the SATs]. The observed gains on the program that looks best should be considerably larger; such gains, however, are evidently extremely rare, either across the board or for specific efforts.

Finally, knowing that a program "works" (gets results) is not the same as knowing what makes the program work. To develop this knowledge, implementation analysis is required. Such analysis is inevitably dependent in some measure on qualitative assessment. Even knowing what made an experimental program work does not necessarily mean that it will work in other places at different times, or that one knows how to make it work in other settings. Here, judgment, systematic thinking and ongoing assessment of both the soft and hard varieties come into play.

In short, just as determining the "essential integrity" of a model takes judgment, so does determining what constitutes credible evidence. Credible evidence of program effectiveness--the best available information--may (and ideally will) include experimental evidence. But, almost inevitably, that evidence will need to be supplemented by other forms of systematic evaluation of program impact. Impact analysis in turn should be complemented by implementational analysis that draws on a combination of observation, interviews, program data and synthesis of related research to try to determine what elements of a model are essential to its success.

Credible evidence will rarely if ever prove that the particular model is superior to all other alternatives. What it must do, however, is make a well-reasoned case that:

- No matter what the results of the program are it is better than no program at all.

- The results of the program are in the desired direction (related to program objectives) and are not only statistically but also socially significant.

- The program is as good or better than any feasible known alternative appropriate to achieving the same purposes.

Finally, it is worth stressing that credible evidence is not acquired at one final point. Rather, it develops, gets refined and needs to be reconfirmed through every stage of the replication process that, in its fullest expression, includes the following steps:

1. Analysis of the problem to be solved;
2. Research and field review of programming approaches that have "face validity" for solving the problem;
3. Specification, development and pilot testing of one or more "models" in no more than a few sites;
4. Refinement of the model based on pilot tests;
5. Broader demonstration testing of the successful pilot model(s)--as refined--to revalidate and assess applicability in a range of locations;
6. A reproduction phase in which one or more waves of replication are conducted in a larger number of sites; and
7. Institutionalization and further dissemination.

Experimental research designs are most likely to be a realistic option in phases 3 and 5 of this rather classical sequence, and the results of that research can be used to establish "benchmarks" for evaluation of replication sites, as several of cases we will review have done.

REPLICATIONS: SOME EXAMPLES

With variations, a multistage, multiyear, common sense process has characterized all the examples of replication we will now review. The following sections report, in part, how these programs developed and used "credible knowledge" but, more particularly, what strategies and "technology" they used in implementing replications.

Franchising in the Private Sector

The most successful and extensive instances of replication do not come from the public sector. They come from the private sector. In all industries, there is a desire to differentiate product--loosely analogous to the desire to find a better program model in the public sector. But, once so differentiated (whether in substance or by packaging) and once they have a product that "works," some industries replicate with zeal. Replication in service industries is the closest approximation to replication of public services, and the example par excellence is fast foods.

Franchising is one approach to private sector replication, and the process of creating strong franchises is essential to the success of chains like Burger King. It begins with a rigorous process of selecting the franchise owner. The process is designed to make sure that the owner has the ability to run a Burger King, fully understands what he or she is getting into and, further, is enthusiastic about and deeply loyal to Burger King.

As described to me by Morris Kurtz, the owner of two Burger Kings in New Jersey, the prospective owner applies to Burger King and, if he passes an initial interview, enters a highly structured 80-hour program, which includes tests of personality, business knowledge and the ability to get along with people (customers and employees). In addition, there is a 10-day intensive course at "Whopper University" covering technical aspects of the business and Burger King operating procedures. "Graduates" of Whopper U. are offered a franchise at a site selected by the company because it meets stringent market criteria. Willingness to relocate is part of the price of ownership.

Thereafter, training, technical assistance and marketing services, as well as strong quality control, are provided by Burger King (the "replicating agent.") Owners will lose their franchise for stealing, failing to meet sales and profit goals, selling unauthorized products, or running an outlet that does not meet cleanliness or other standards imposed by periodic visits from corporate inspectors. Some of the training and support is quite predictable: there are seminars for owners on changes in the product line, customer relations and other aspects of running the business; there is a training package for new employees. But some of it is surprising. Employee retention is a major concern right now, according to Kurtz, who notes that "lots of my employees are 16- and 17-year-old kids. These are their first jobs and you really have to be able to relate to them. Some have lots of problems, and they talk about them at work. I recently took a course on counseling teenagers so I could do this better."

Kurtz finds all the training "terrifically helpful," as he does the highly detailed manuals and operating rules that govern every aspect of the operation. "The key," he says approvingly, "is that it's so simplistic . . . The bottom line is consistency."

Food service and human services are, of course, very different industries with very different types of workers. But they share some generic elements when it comes to the process of replication, and these make for interesting comparisons. These include selection of sites where the product will sell (program will work) and the selection and screening of owners (program managers), the cultivation of loyalty in owners (managers); the process of training, technical assistance and support; and ongoing evaluation and quality control. These functions, as we shall see, also appear integral to successful replications in the public sector. Striking, however, is the tremendous amount of effort and investment that goes into all these aspects of implementation in the fast food chains--a level rarely contemplated, much less achieved, in the public sector. If selling hamburgers takes this kind of rigorous implementation, can one expect less will succeed in providing education or training?

The differences, however, are as important as the similarities. Burger King operates with a very strong and clear incentive: profit. As long as profits follow from the Burger King method, owners have no reason to resist it. Burger King also has the ability to get rid of poor owners with relative ease and, through its franchising agreement, a strong and clear line of authority.

Incentives, like goals, are less clear and more diffuse in the public sector. Further, authority relations are weak, hemmed in by everything from political to civil service to structural constraints--especially in the "loosely coupled" education system (Weick, 1976). But all this means is that attention to incentives and the proper use of authority become particularly critical in public program replications--another topic to which we shall return.

Franchising is perhaps the best example of one polar approach to replication: a highly specified, high-fidelity/low-adaptation approach. It represents one end of a continuum along which to array and explore the examples of public sector replication, and it suggests some of the generic issues that they and any planned replication must address.

Replication in the Public Sector

I will briefly describe five examples of fairly large-scale replications in the public sector. These range from fairly "closed" efforts aimed at reproducing well-specified program models to fairly "open" efforts designed to engage practitioners in adapting key principles to their local situations. The former type, closer to the franchise approach, I will call "product-oriented" replications; and the latter "process-oriented." (This is not to imply the former lack process or the latter are empty of product or program.) After briefly presenting these cases, a number of the generic tasks and issues they illuminate will be discussed in Chapter IV.

The Comprehensive Competencies Program (CCP)

One of the strongest findings to emerge from the frenzied "knowledge development agenda" of YEDPA was that basic academic skills are a strong predictor not only of youth labor market success, but also of success in training programs (Taggart, 1981; Sum, Taggart and Berlin, 1986). (Sometimes the greatest utility of research is that it puts the obvious in boldface.) Further, research on the Job Corps and some other studies suggested that competency-based, computer-managed basic skills programs were highly effective with out-of-school, disadvantaged youth. Drawing on this research, The Ford Foundation in 1981 commissioned Robert Taggart, former head of DOL's Office of Youth Programs, to develop a competency-based basic skills program that would fully use technology and pull into one system "the best" of available software and curriculum appropriate for this target group.

CCP's curricula cover "basic, intermediate and advanced" math, reading and other academic skills, English as a second language, occupational information, "life skills" and more. The core of CCP is an elaborate computerized management system that tracks student use and progress. It provides instant scoring of tests, directs students to the appropriate print or software materials for the next lesson, and provides extensive management information about the entire program. CCP is billed as an "individualized, self-paced, competency-based, open-entry/open- exit, multimedia" approach to basic skills instruction. While students work at their own pace, the teacher in a CCP program, freed from leading the class and from paperwork, can assist students on a "one-to-one" basis.

With Ford support, Taggart's Remediation Training Institute (RTI) tested the program in a number of community-based organizations and reported large test score gains in a very short time. Taggart (1987) claims that the average CCP student gains 1.0 grade level in reading and 1.4 grade levels in math after 28 hours of participation in the program--at a cost of $250 per grade level gained.

Bolstered by results of the pilot test, and supported by a combination of grants and a $1.5 million revolving loan fund from Ford to purchase computer equipment, RTI formed a nonprofit subsidiary, U.S. BASICS, to market and disseminate the program. CCP's replication strategy is franchising. Unlike Burger King, U.S. BASICS does not make a profit (though the organization hopes to become self-sustaining), and it is not restrictive about the use of other materials in a CCP center. Otherwise, however, the process is strikingly similar.

Community-based organizations, schools and others interested in using CCP submit an application, visit a CCP center, receive a site review by a CCP-certified technical assistance provider, submit a basic plan to U.S. BASICS and, if accepted, they enter into a one-year CCP Use Agreement. The agreement spells out in considerable detail what the user must do to maintain the franchise.

Requirements include purchase of hardware and core software from U.S. BASICS, payment of a $1,000 user fee, participation at CCP training sessions (paid for by the user), operation in compliance with the user's approved plan, and extensive reporting. Reports are used by U.S. BASICS to monitor performance of franchises and their compliance with operating plans and the user agreement. Unsatisfactory performance can and has led to termination of some agreements. Each year, a new plan must be submitted and a new user agreement is signed. U.S. BASICS, for its part, provides continuously updated curriculum material, system user support, a newsletter, and ongoing training and technical assistance.

Marketing is critical to replication of CCP, since sites must purchase the franchise, and U.S. BASICS has developed a marketing staff. One service the company offers is technical assistance in fundraising to users who cannot support the program from existing budgets.

As a replication strategy, U.S. BASICS has been highly successful. CCP is now being used in more than 300 sites, serving some 40,000 students per year, most of them young adults. The biggest markets have been community-based organizations, alternative schools, and job training programs. Typically, CCP is used as the education component of a training program. According to U.S. BASICS staff, the early gains reported by CCP have held up, though there is (as one would expect) considerable variation around the mean.

CCP is not without its critics. Many computer-assisted remedial programs show quick grade-level gains, but some researchers believe these are largely the result of participants simply brushing the rust from skills they had acquired previously. Some criticize competency-based learning in general on grounds that it is too narrow, or that learners are unable to transfer skills to new settings and problems. There has been no controlled, third-party study of CCP's effectiveness, and some of its literature includes "research findings" that smack more of marketing than of social science.

These questions aside, CCP has clearly identified a market and developed a replication strategy that in a few years has had great success in penetrating the market. The program was consciously designed to be all but "teacher-proof," in recognition of the limited teaching backgrounds and capabilities of educational staff in community-based organizations. This helps explain why U.S. BASICS has encountered some resistance in its efforts to market to school systems[9] (with the significant exception of alternative schools), but great success with community-based organizations whose clients typically have an overwhelming need for remedial education. Opportunities Industrialization Centers (OIC) is one of the biggest CCP users (34 of 72 OIC affiliates). OIC's Ben Lattimore says:

> CCP's been great for us. It gave us a means to standardize programs, provide a sequence of offerings and maintain quality control. Before, we had a rag-tag air force of individual pilots flying their own planes. . .

> There's been great acceptance. After all, they [U.S. BASICS] come in with a bag of goodies, a pre-funded computer, programs, training, and great ongoing support. Most idiots would accept that.[10]

Summer Training and Education Program (STEP)

While still in early stages, P/PV's STEP program, which was operating in 50 communities serving over 5,000 students in 1989-90, is already a large planned replication. It is one of the few to follow all the steps of the classic R & D process--from experimental pilots through demonstrations to replication. It is also a "product-oriented" replication, requiring high fidelity to the model and, in the words of a P/PV vice president, its replication strategy is "toward the Burger King end of the continuum." (Some aspects of STEP, however, are loosely specified, more process-oriented and rely considerably on the judgment of local practitioners.)

Like CCP, STEP evolved from earlier research and had initial Ford Foundation support. Ford's Gordon Berlin had been struck by several studies suggesting that, unlike their more advantaged counterparts, disadvantaged youth lose ground academically over the summer, when school is out. At the same time, YEDPA research (and other studies) had concluded that summer job programs do little or nothing to improve students' employability skills. P/PV was funded by Ford to develop a program whose most basic goals were to combine summer remedial education and work experience for 14- and 15-year-old disadvantaged youth as a way to halt summer learning loss and add value to the summer jobs experience. In addition, the program was designed to incorporate a "life skills" curriculum, with the goal of influencing teenage sexual behavior and reducing teen pregnancy, a strong correlate of dropping out of school for teenage boys as well as girls. The ultimate goal of STEP, in fact, is to provide a series of educational and personal growth experiences that prevent dropping out.

P/PV designed the basic STEP program in early 1984. It was pilot tested in several sites in the summer of 1984. Initially, P/PV asked sites to develop their own remediation curriculum. But this proved to be a difficult and unpopular task at the sites, so P/PV developed its own.[11]

The program was designed as a 15-month experience, combining two summers and the intervening school year (P/PV, October 1985). Major program elements include:

- Ninety hours of basic skills remediation designed by P/PV, in an "individualized and self-paced" curriculum;

- An 18-hour life skills curriculum, also designed by P/PV, which covers a broad range of topics (decision-making, employment information, workplace behavior) but is primarily aimed at developing responsible social and sexual attitudes and behavior;

- Work experience, provided through the federally funded Summer Youth Employment Program (SYEP) in minimum wage jobs in such areas as clerical, maintenance, parks and recreation work;

- A school-year support component, including such services as counseling and tutoring, to help participants sustain any academic gains achieved during the summer and successfully negotiate through the school year.

Implementing this model requires cooperation between school systems and JTPA job training agencies in each site, as well as a program management agency and all agencies responsible for providing worksites, delivering remediation and operating the year-round program. There is considerable variety from site to site in the agencies that perform these tasks, though typically a Private Industry Council, a school district and/or a nonprofit youth agency are involved. Thus, to its educational program components, the STEP model adds the complexity of interagency cooperation.

Research on the 1984 pilot program showed STEP's operational feasibility and provided solid evidence of tangible impact on basic academic skills and participant knowledge of contraception and the burdens of teenage parenting. It also disclosed some areas--notably, the design and delivery of remediation--where more work needed to be done to assure smooth and effective implementation. Based on the generally encouraging findings, P/PV moved from the pilot stage into a five-site demonstration, supported by Ford and other private funders. The demonstration, which included a strict random assignment experimental research design, continued to produce strong results (Public/Private Ventures, 1988). Accordingly, the STEP program moved into a first wave of replication, adding 11 sites in 1988 and 33 additional ones in 1989. With 100 participants in most sites (and more in some) STEP served over 5,500 youth in 1989.

The U.S. Department of Labor and participating state governments have added resources to support replication. Preliminary research showed that the first wave of replication sites not only matched but on average exceeded the gains achieved in the demonstration.[12] The basic elements of the replication strategy evolved by P/PV include:

- Involvement of state government in selecting and supporting demonstration sites--a strategy that P/PV hopes will build institutional support for STEP after its involvement ends;

- Intensive technical assistance and training, provided by P/PV, with primary focus on planning the initial summer program;

- No control groups, but site collection of data on participant outcomes (primarily pre- and posttest scores) to enable P/PV to compare results of benchmarks established during the demonstration and to support local data needs;

- Insistence on "high-fidelity" replication of the model's basic parameters (15-months, four service components, public schools/employment training agency cooperation, and basic management functions);

- High-fidelity use of P/PV's life skills and remediation curricula;

- Flexibility in the nature of summer jobs provided; and

- Considerable reliance on local practitioners to design and develop the school year support component.

As noted earlier, P/PV has followed a classic model of development, testing and dissemination through hands-on replication. But note as well that, for pragmatic reasons, research elements of this model have been somewhat truncated. While early findings were unusually powerful in magnitude and "clean" in research design, the long-term impact of STEP remains unknown. Still awaiting longitudinal research are the answers to two key questions: Are the effects of an intervention at ages 14 and 15 sufficient to reduce the dropout rate of STEP participants? Do increased knowledge of birth control and the consequences of early pregnancy reduce teen pregnancy rates?[13]

Jobs for America's Graduates (JAG)

JAG's origin lies in Jobs for Delaware Graduates (JDG), a program developed in 1978 with support from YEDPA. JDG in turn incorporated many of the program elements of 70001 Ltd., a preemployment and education program for out-of-school youth previously run by JAG's president. 70001 itself has been successfully replicated in over 100 communities.

Quasi-experimental research a decade ago showed that 70001 helps participants get jobs more quickly than comparison youth, and that their employment was somewhat more stable. However, the benefits to participants eroded after a year, apparently because participants lacked the basic academic skills to maintain their initial advantage over comparison youth (Public/Private Ventures, May 1983). (70001 has made changes in its program model to strengthen educational offerings.)

In effect, JAG bypassed this problem by initially concentrating its program on high school seniors--i.e., students whose basic skills are good enough to graduate from school. Early nonexperimental studies of JDG showed generally positive results. Subsequent quasi-experimental research has shown that JAG participants are more likely be employed nine months after graduation than comparable (but not randomly selected) youth, and that they earn between $800 and $1,000 more in their first year after graduation than nonparticipants (Sum, 1987).

True to its origins, JAG remains a statewide program, providing employability services and job placement assistance in senior high schools. There are now JAG affiliates in 14 states, serving over 18,000 youth per year in over 200 high schools.

JAG provides a combination of career guidance, motivational student activities, and preemployment training to juniors and seniors, as well as job placement assistance to graduates. There is a national JAG board that contracts with a company run by Kenneth Smith, JAG's founder and president, for management services, including provision of curriculum materials, training and technical assistance to local programs. The primary functions of the national organization are to market the program, help set up new sites, and provide ongoing monitoring of local programs. With a few exceptions, management of the actual program is in the hands of state affiliates of JAG, such as Jobs for Tennessee Graduates, Jobs for Ohio Graduates, etc., which, in effect, are state-level franchises. "What we are marketing," says Smith, "is a statewide school-to-work transition program."

JAG's replication strategy is unique both in its use of statewide organizations and in its political astuteness. Its first chairman was Governor DuPont, and the JAG board has always read like a "Who's Who" in national politics, with both parties well represented. Its political weight is supplemented by CEOs of major corporations and heads of national advocacy organizations. JAG relies on "heavyweight" board members, at both national and state levels, to sell the program and to give it staying power. In its first wave of replications, JAG entered four states, and Smith believes that:

> Three-quarters of our ability to sell the program came from who we had selling it--one governor to another--with the research quite secondary. In fact, at the beginning, we did not have much research to go on. That came later, and it has become more important--particularly in getting credibility for the program with middle managers who have to support it after the governor said yes. At the political level, research was icing on the cake. At the staff level, it's more likely to be the cake.[14]

It also helped that JAG initially was armed with $3 million a year for three years from federal and private sources, including corporations and The Rockefeller Foundation. These funds covered about 50 percent of the cost of each state program, and JAG is now in the process of weaning those states to local funding. The remaining 10 states, however, have paid the full cost of program implementation, and JAG recently began to charge each state affiliate a service fee--ranging from $5,000 to $50,000 per year, depending on the size of the state program. These fees provide some relief from the need to raise private dollars for JAG's $600,000 central budget.

Entering a state, JAG seeks a "champion" in state government--usually either the governor or the chief state school officer. It then builds a board that, in addition to these officials, includes three or four CEOs, labor and civic leaders, the state Chamber of Commerce head and, where laws permit, a key legislator. It typically takes two years from initial contact to the start of a program, as the board is developed, funding is secured and staff are hired and trained. The boards are responsible for overseeing the growth and the implementation of the program in each state, and they provide continuity after turnover of elected officials. In Smith's words, "We've survived turnover in many states and won hundreds of budget fights because of board support." In some states, the boards have successfully pushed for separate legislative appropriations, thus institutionalizing the program in the state budget.

JAG permits states to innovate and augment the model. It pays at least lip service to the idea that every state is different. But those differences are more likely to affect incidental aspects of implementation than the essential program. In fact, JAG has a list of "non-negotiable" elements to which each affiliate must adhere. These elements, which assure that JAG looks pretty much the same from state to state and school to school, include:

- Creation of a nonprofit organization, with involvement of JAG in its formation;

- JAG involvement in initial budgeting and funding decisions, as well as in hiring key staff;

- Staff training by JAG;

- Creation of a student organization for motivational activities;

- Use of the basic JAG program model;

- Use of the JAG preemployment curriculum (unless the state can convince JAG that it has a better one--and none has).

- Reporting of data on program performance to JAG; and
- Submission to an annual review by JAG for purposes of "re-certification" as a JAG affiliate.[15]

By and large, this quasi-franchising approach appears to have worked well. Only one state has dropped out of JAG. Twice, JAG began with a city organization instead of statewide program. In one instance, the program then went statewide; in another, it did not. States do vary in the extent of intrastate replication. For example, Tennessee has 90 high schools participating, while Pennsylvania, still in its early stages, has 155 students participating. JAG has been adding states at the rate of two a year for five years, and expects to continue to do so.

Experience Based Career Education (EBCE)

Unlike our other examples, EBCE was a replication initiated by the federal government--specifically, the Education office of what was then Health, Education and Welfare. And unlike the Head Start and Follow Through Planned Variations, EBCE had a long run and substantial success, as indicated both by research results and its spread from concept in 1971 to programs in some 200 different school districts (approximately 1% of school systems) by 1979.

EBCE differed from the replications described above (and the Head Start and Follow Through Planned Variations that began shortly before it), in that it represented less a program model than a vision of education. That vision was grounded in the philosophy of a seminal thinker, John Dewey, and emphasized learning that was experiential, hands-on, individualized and competency-based --as opposed to prevailing modes of classroom instruction. It incorporated aspects of a number of strong educational currents of the late 1960s: the concept of "schools without walls," which viewed the community as a learning laboratory (Silberman, 1971); the career education movement (one of whose leaders, Sidney Marland, was then U.S. Commissioner of Education (Marland, 1974); and growing criticism of traditional models of vocational education, which were variously argued to be ineffective, counterproductive or too narrow in training students for specific jobs in a fast-changing world (Grubb and Lazerson, 1975).

In fact, EBCE was four versions of this vision--four broad models with some common themes--which in turn were "process-oriented" replications. Each permitted and encouraged local school districts to apply creativity in adapting the models to their situations. The National Institute of Education (NIE), which planned EBCE as a replication from the outset, asked four of its regional laboratories to develop models that would share a number of common characteristics. All were to achieve a balance of career and academic learning for a cross-section of high school students, through a program that included:

- An organizational structure for involving people and facilities in the community as student learning resources;
- A method for identifying and analyzing community learning experiences;
- A method for assessing student interests and abilities to guide individualized program planning;
- Combination of the role of teacher and counselor;
- Activities that permitted students to explore different career areas and clarify career choices; and
- Guidance and activities to develop competencies needed for dealing with problems and decisions of adult life. (NIE, 1976)

In 1971-72, the four "labs" developed their programs. In the next three years, they were pilot-tested in four schools. The pilot programs were conducted as experiments. Students who expressed interest in EBCE were randomly assigned to treatment and control status, and the program was evaluated by the Educational Testing Service.

There were considerable differences in the characteristics of the models. For example, Research for Better Schools in Philadelphia developed EBCE primarily as an adjunct to the regular school program, while the Northwest Regional Education Laboratory in Portland, Oregon developed an alternative education model in which virtually all learning--academic, career and life skills--took place in the community, supplemented by student journals, self-paced instructional materials, and individual or small group counseling. The Educational Testing Service study (Watkins and Corder, 1977)--and a number of other evaluations summarized in Owens (1982)--consistently found positive (if modest) results on common measures that spanned all four models, especially in areas relating to career knowledge, self-concept and enjoyment of school. Findings on basic skills ranged from "no difference" to slightly better for participants. Clearly, EBCE was not a dramatic improvement over "regular" schools, but the evidence established the approach as having modest benefits and viability as an alternative approach to education.

In 1975-76, NIE funded the labs to prepare materials and technical assistance for expansion. The program spread to an additional 14 schools, showing in the process that it could be operated at costs roughly comparable to the regular school program.

Results continued to be positive (Owens, 1982), and NIE then launched an even broader replication. It disseminated information widely and held a series of regional conferences at which the labs presented their approaches to local school officials. NIE provided funds, through Title IV of the federal education act, to school districts to purchase training and technical assistance from the regional labs.

In the next two years, EBCE spread rapidly to at least 190 districts in 48 states (Spotts and Hampson, 1979), with the typical program serving about 100 students. The availability of federal funds no doubt spurred this rapid expansion. Indeed, one replicator felt that some sites picked up the program simply because funds were available.[16] But an important contributing factor appears to have been that the program found a market among educators and, particularly, teachers who were dissatisfied with regular school programs and sought an outlet for their creativity and dissatisfaction (Farrar, De Sanctis and Cohen, 1979).

The approach to replication varied among the regional labs, but it generally emphasized a process of local decision-making, with technical assistance, training, operating manuals and optional curricular materials provided by the laboratories. Labs had neither fiscal nor reporting control over sites, and the technical assistance was modest. The Northwest Lab, for example, attempted to keep the cost of its TA package down to about $5,000 per district. This, according to Rex Hagans, then project director at the lab, kept costs in a range the districts (many of which were small) could afford.[17]

A Huron Institute study of this replication effort (Farrar, De Sanctis and Cohen, 1979 and 1980) concluded that local school officials were like "guests at a garden party" who brought widely different ideas about EBCE to the replication, producing adaptations that ranged from "fine-tuning to wholesale restructuring" of the lab models. The Huron study seems to assume that adaptation undermined the goal of replication. But many of the examples they cite, such as teachers who emphasized counseling more than instruction, seem entirely consistent with the aims of the replicators. Hagan points out that the essence of the Northwest Lab model was its shift in the locus of schooling from classroom to community, and its trinity of education goals: academic, career and life skills. The Northwest Lab encouraged local adaptation. Its training programs stressed the kinds of functions EBCE staff would have to perform, but did not prescribe a specific staffing pattern (beyond the guideline of four teachers for 80 students). It provided curriculum materials and operating manuals, but viewed them as optional, encouraging schools to adapt to increase local fit. In Hagan's view, the fact that most of the 50-plus sites following the Northwest Lab model did in fact build their programs around its basic principles--itself a dramatic departure from traditional schooling--meant that the "essential integrity" of the model was preserved in most replication sites.[18]

From 1971 to 1979, NIE stayed with EBCE, carrying the model from concept through experimental evaluation of pilots to broad-scale replication. This was uncharacteristic continuity on the part of a federal agency. Unfortunately, no systematic evaluations of the replication sites were carried out.

In 1979, NIE lost interest in the program--not because of any negative evidence, but because of changing priorities and turnover in staff who had been champions of the model. The regional labs were no longer funded to work on EBCE, and there were no longer special funds set aside for school systems to initiate the program. Replication activity withered away, and so did many of the sites. While some EBCE programs are still operating, no one knows exactly how many.

Coalition of Essential Schools

In 1979, Theodore Sizer and a number of colleagues, supported by six foundations, conducted a survey of American high schools under the sponsorship of the National Association of Independent Schools and the National Association of Secondary School Principals.

Their primary method was radically simple: they visited schools to observe and interview students, teachers and administrators. They came away with some overriding conclusions: the schools were drearily similar from town to town and state to state; they were highly impersonal, to students and teachers alike; most teaching consisted of lectures, and most students were passive and bored in the classroom; teachers were burdened with heavy-handed administration, too much paperwork, competing central office initiatives to implement, and too many students. Sizer reported that the school program resembled a "shopping mall," lacking philosophical and educational coherence, and everywhere there was, "paralysis of imagination."

Not content with mere recording or analysis of the survey (Sizer, 1984; Powell, Farrar and Cohen, 1985), Sizer, former dean of the Harvard School of Education, former Andover headmaster, and now a professor at Brown University, decided to do something about it. He developed what has become known as the Coalition of Essential Schools (CES).

Initiated in 1984, the Coalition represents an interesting combination of "bottom-up" change, school restructuring, and a highly adaptive, process-oriented approach to replication. Like EBCE, this replication depends more on a vision of schooling than it does a model. Unlike EBCE, its aim is not simply to provide an alternative education for some students in a school, but to transform the entire institution. Sizer calls it "a coalition driven by a set of ideas: some ideological, some specific, some grounded in what have long been recognized as basic truths about education, but most of it just common sense. . . . A school is essential if on its own terms it is trying to organize itself around these ideas."[19]

Sizer describes his replication process ("replication" is my word, not his) as "an effort to sustain a conversation about some important ideas in the faith that people will then learn how to adapt them effectively in their own schools."

The process has now brought 55 schools into the Coalition (a number expected to double in the near future). It is much less structured than that of other replications I have reviewed, but there is more to it than "conversation." And, while not tightly specified, the Coalition's guiding principles suggest a model writ large, with lots of room for school-by-school creativity and variation.

CES has developed nine guiding principles for an Essential School, as well as a contractual set of obligations that bind the Coalition together. These cover governance (Boards of Education must agree to delegate broad authority to Essential Schools for four years, teachers must be involved in curriculum decision-making and planning); school organization (no teacher responsible for teaching more than 80 students, teachers should adopt multiple roles--e.g., teacher, counselor, learning "coach"--and a budget not more than 10 percent above the school district norm); curriculum (the curriculum should be stripped to essential intellectual competencies); and instruction (students should be engaged, active learners in inquiry-based classrooms; diplomas should be awarded for exhibited competency, not accumulated credits).

In short, while the Coalition is an effort at restructuring, it is restructuring to achieve an educational vision grounded in a set of principles that range from the general to the specific (Brandt, 1988; Coalition of Essential Schools, 1988). It thus attempts to avoid the trap that (I believe) awaits many of the highly publicized restructuring efforts around the country, which have concentrated on changes in governance (e.g., Dade County), or changes in staffing patterns and teacher roles (e.g., Rochester) without having first addressed the substantive educational question: restructuring for what kind of education? (For an early report on restructuring that suggests that the primary outcomes so far have been a shift in power--from middle management to the schools--see Elmore, 1989.)

To develop and spread the concept of Essential Schools, Sizer has an eight-person professional staff at Brown. Another 10 to 15 people serve as state coordinators in the five states (with two about to be added). Their Departments of Education have come on board in a joint effort to expand CES in partnership with the Education Commission of the States (ECS). Also, there are about 20 teachers from participating schools who form a "Coalition faculty," and these teachers, along with Coalition staff, bear major responsibilities for running workshops at state and national meetings of the Coalition. (The Coalition faculty reflects Sizer's conviction that "the most credible folks to other teachers are the people who are doing it.")

Workshops are augmented by videotapes of instruction in Essential Schools. While the Coalition forswears packaged curriculum or operations manuals, just as it abhors use of the word "model," it has developed model budgets to help schools figure out how to redeploy staff to progress toward the 80-pupil limit for any subject matter teacher--no small exercise in urban school systems where the number is often in excess of 150. In addition, the Coalition publishes a journal, a newsletter and a number of articles discussing aspects of restructuring.

When schools express an interest in joining the Coalition, they are usually advised to wait for as much as a year until they have attended some Coalition workshops, fully understood the concept and what it will require of them, and gathered support from a strong majority of the faculty. Once a school is admitted, it receives $50,000 a year from a combination of state funds and corporate contributions to pay for developmental costs.

Consistent with his overall philosophy, Sizer believes the success of Coalition schools must be judged largely on the basis of criteria developed by each school. He has eschewed efforts to quantify success, believing such efforts would be premature in a young initiative that envisions a five-year change process in each school. However, the Coalition has commissioned a distinguished group of visitors to observe schools and assess progress, coordinated by Gerald Grant, professor of education at Syracuse University. Based on their observations and other sources of information, Sizer believes that "about 33 percent have done nothing, either because they are paralyzed or have changed leaders; about 33 percent are starting the conversation, but there's nothing yet to look at; and another 33 percent are doing something, with six or seven schools really having something to crow about."[20]

Sizer is pleased that the most successful schools are in urban areas and serve low-income students--a fact he attributes to the inability of such schools to argue that what they are currently doing works. In the long run, he declares, "if one out of five schools make it, that's terrific."

IV. LESSONS FROM EXPERIENCE: THEMES AND VARIATIONS OF REPLICATION

The five cases discussed here suggest the wide variety of replication approaches, and each represents--with some significant qualifications--a successful planned replication. That is, each represents an effort to develop program models or approaches from a systematic consideration of social science evidence, broadly understood. "Credible knowledge" underlies all these efforts, and each (with the exception of the Coalition of Essential Schools) began with pilot and/or demonstration activities that were evaluated (using experimental or quasi-experimental designs in three instances) and found effective. Each exhibits a fairly clear sense of the "essential integrity" of the model, ranging from CCP's tightly defined and tightly controlled franchise approach to the more open-ended but principle-bounded examples of EBCE and the Coalition. Each succeeded in creating a substantial number of reproductions of the original model that, within the terms of the endeavor, exhibited that essential integrity.

There are important qualifications to this claim of success. Only EBCE and STEP used experimental research designs. Only EBCE and CCP had all the evidence in before moving to replicate. Critics may quarrel with methodology or limits to the research design in every case. It is too early to say much about the impact of CES, and skeptics may be bothered by the lack of interest in hard data within it. The jury is out (or, in the case of EBCE, the data are lacking) on the long-range programmatic impact and institutionalization of all these examples. Will they spread within locales as well as across locales? (STEP and JAG have shown some propensity to do so; EBCE apparently did not). Will they endure after the replicating agency and its technical assistance move on? Will there be sufficient sites to constitute a critical mass--thus beginning to influence basic paradigms of program planners outside the planned replication, or seeping into broader practice in ways that, like research itself, are hard to trace?[21]

Qualifications aside, these examples constitute prima facie evidence that planned replication can be managed, can provide continuity over long periods of time, can be effective in reproducing programs based on credible evidence, and can maintain solid levels of program performance.

The critical questions thus become:

- What are the key factors in successful planned replications?
- Given a continuum of choices for replication strategy (from high-fidelity, highly specified "product-oriented" replications to high-adaptability, vision- or principle-driven

"process-oriented" replications), are there guidelines helpful in determining what situations require which strategic choice?

- What are key technical issues for replication?

- For what set of problems is replication a potential answer?

KEY SUCCESS FACTORS

Despite considerable differences in strategy, the five planned replications reviewed here suggest some elements common to successful replication. Given limits of the data and paucity of cases, I hypothesize that the following are critical success factors:

1. Finding a Market

The first, and perhaps most important factor, is that successful planned replication appears to be inherently entrepreneurial and market-driven, as opposed to a bureaucratic and legislatively driven. Interest in applying social science to discovering models that work in the late 1960s and early 1970s often implied a top-down process of social/political engineering. Federal research would produce the answer and federal initiative would spread the practice to every nook and cranny of the land.

Rather, successful replications thrive because they find a market. The market may be stimulated by the availability of federal funds, as with CCP and STEP, or by federal programming initiatives, as with EBCE and NDN, but the product is sold, not mandated. Some look for a small niche, others for a potentially large share, but each of the cases we have reviewed had a market: CCP has the "high need/limited capacity" educational segment of the employment and training business; STEP has the professional and programmatically weak world of summer jobs; JAG has the nearly empty zone of school-to-work transition services; EBCE has the teachers and administrators who are dissatisfied with "standard brand" education, as well as the market created by availability of federal funds for the replication. CES's market is any school attracted to its vision.

2. Incentives

Viewing planned replication as a market-driven activity casts in bold relief the issues of incentives. In the private sector, many incentives operate, but the dominant incentive is profit. This incentive is strongest for owners, but it is frequently reinforced in the formal compensation plans of executives and others. Further, owners and managers have considerable freedom and authority over such matters as capital investment, products or services offered, and personnel to deploy in pursuit of profit. This strong incentive/authority field supports replication of franchises or chains in the private sector, just as the promise of ultimate rewards and a strong hierarchy have historically supported the enormously successful replication efforts of religious organizations, such as the Roman Catholic Church.

In the public sector, program results are the closest analog to profits, but the analogy is a poor one. For in the public sector, there is often legitimate disagreement over what constitutes a desired result, over the importance of a result, and over the most appropriate measure of results, as well as technical difficulty in measurement. Public managers rarely approach the degree of authority and control over investments, product mix and personnel enjoyed by their private sector counterparts--constraints that diminish the willingness to take risks. When results are good, there is no distribution of profits and there may not be rewards of any kind. When results are bad, there are often no consequences. Incentives without consequences are, of course, a contradiction in terms.

This does not mean that the public sector lacks incentives. Rather, it means the incentives are often complex, diffuse, variable, and hard to implement. Further, while there are incentives to get results (and plenty of public officials who are intrinsically motivated to get them), there are also incentives to do nothing. In fact, many incentives in the public sector are double-edged when it comes to replication. For example, the desire to build "turf"--which approximates power and prestige--could lead an administrator to opt for a replication that represents a budgetary expansion. But, if it requires diverting funds from existing programs or sharing power with a collaborating entity, it could lead--as it often does--to the opposite decision.

Any attempt at an anatomy of the incentives affecting replication lies beyond the scope of this paper. Indeed, a definitive anatomy is probably impossible, because public sector incentives are, as the example of turf control illustrates, highly situational, and thus not subject to easy formulation. However, it is self-evident that replications will fail if they are not consonant with the local incentive system, and there are a few key incentive issues to which replicators must pay special attention. CCP and STEP are both relatively straight forward responses to incentives derived from federal legislation. But the fit is not always so easy.

First, because incentives are diffuse and authority is weak, it is often necessary to sell "up and down" the system. Convincing an elected official or a superintendent to buy into a replication is rarely sufficient. Senior managers, middle managers, teachers or other program staff also have to buy in, and what is attractive and effective to one level may not be meaningful to another.

An example is Ken Smith's observation that peer endorsement of JAG was often what sold governors, while research findings were more important to middle managers. In "process-oriented" replications like CES, the highly adaptive, participatory process of defining needs and planning solutions is a strategy for trying to ensure a good fit. There are indications that CES is experiencing problems in some sites because it has not paid enough attention to selling multiple levels of the institution (Timar, 1989). In STEP, the availability of a proven model that meets JTPA regulatory requirements to offer remedial education as a component of the summer jobs program appeals to many JTPA administrators. But it is the quality of the curriculum and the technical assistance, P/PV staff believe, that sells program staff and teachers.

Career advancement is a powerful incentive in the public sector, especially for managers and elected officials. Innovation is frequently associated with advancement, and when replications are packaged as innovations--as all the cases reviewed here clearly have been--the incentive of career advancement may come into play, as it often does in successful change efforts (Huberman and Miles, 1984). However, this too is a double-edged incentive. In systems with few tangible incentives for success, pride of authorship may be its own reward. As Anthony Downs has written, the public sector is peculiarly cursed by the "Law of Compulsive Innovation." Upwardly mobile bureaucrats move up by making a name for themselves as innovators, not by sticking around to make someone else's program work--or even by waiting long enough to make their own program work. Adaptability, from matters as small as name changes (Jobs for X's Graduates) to significant variations in program design, helps package a replication as an innovation. The incentive of career advancement can also be tapped by the prestige of organizations associated with the replication. JAG enjoyed White House support and has had continued backing of national foundations, as have STEP, CCP and CES.

Whether they are attracted by evidence of effectiveness, by the desire to be known as an innovator, or by the power and prestige of a program's sponsors, the fundamental point is that effective adoption of a new program does require leadership commitment. The literature on change makes it clear that, while change may require buy-ins from many levels of the organization, it does not happen, even in the limited authority realms of the public sector, without administrative pressure and support (McLaughlin, 1976; Huberman and Miles, 1984). But insofar as successful innovation produces mobility, it also creates a problem for new programs: turnover among key decision-makers and the loss of at least one of their champions. The need for continuity is hard to satisfy in the world of electoral politics, and it is harder still when the persons most attracted to change are among the most mobile. Thus, most of the programs reviewed here have attempted to create multiple bases of support to reduce the dependency on any one individual: e.g., JAG's multi-sector state boards, and both P/PV's and CES's involvement of both state and local officials. One wonders whether EBCE might not have benefited from greater cultivation of states as partners in the enterprise, since they--not NIE or its regional labs--bear the main responsibility for governing, funding and policy-making in education.

Perhaps the trickiest and most important incentive issues, however, involve funding. Funding has two dimensions: who funds the replication, and how expensive is it on an input cost basis. It is not difficult to "buy" replications--that is, to put up the money for a site to replicate a program. It is harder to sell it--to have the site put up the dollars. But in the long run, the harder choice clearly increases the likelihood that a program will be institutionalized; that is, that it will not fade away when the external dollars to support it dry up. In the pilot and demonstration stages, external funding of a model may be required. In replication, it is at best a mixed blessing. Mitchell Sviridoff, drawing on his experience with the Local Initiatives Support Corporation (LISC), a widely replicated strategy for community development, suggests the ideal approach: "LISC won't go into a city unless we're matched, dollar for dollar, up front, with money going to LISC. LISC won't accept promises or in-kind support. Put the money on the table."[22] Similarly, U.S. BASICS is considering limiting new sites to locations that make two-year commitments.

With varying success, the examples reviewed here have all paid considerable attention to local support issues. After initial pilot testing, EBCE was field-tested further in part to make sure it could be run by districts within prevailing per pupil costs. JAG claims its cost-per-placement is half the JTPA average. CCP is very cost-competitive with alternative means of providing basic education. STEP adds $1,600 per student to the cost of summer jobs, but P/PV seeks to have operating costs borne by local funders (and in fact objected, to no avail, when one state Department of Education decided to pick up costs with discretionary funds). And CES tries to limit costs to 10 percent above current per-pupil expenditures.

In theory, the unit cost of a program should not matter if the benefits outweigh the costs. But the investment approach to social programs carries little weight in the face of political incentives to serve the greatest number at the least cost. P/PV's Ventures in Community Improvement (VICI) program, which provides job training to unemployed people while rehabbing houses or making other community improvements, is widely cited in the literature as a "proven" model whose benefits outweigh its cost; but its high initial cost per participant has limited the success of attempts to replicate it. Similarly, Manpower Demonstration Research Corporation (MDRC) massive Supported Work Demonstration, one of the longest, most expensive ($85 million) and best-researched efforts in the annals of social science, concluded that the program was a cost-effective way to move welfare recipients into jobs (MDRC, 1980). Officials of MDRC and the U.S. Department of Labor, including then-Secretary Ray Marshall, began planning a 50-site replication. Apparently, the question of cost per participant was brushed aside until an assistant secretary in charge of public relations asked the question while developing a press release. The answer ($10,000 to $12,000 per participant, in 1980 dollars) helped spell the end of the program.[23]

Successful replications must navigate the shoals of these complex, double-edged incentive systems: turf considerations; the need for different incentives at different levels; the pluses and minuses of career advancement as an incentive to innovate, and funding issues. The list could go on. Replications such as CCP, JAG and STEP, which operate in a niche of low bureaucratic and institutional density, will have an easier time of it than efforts to replicate programs that envision major institutional change or operate within encrusted bureaucracies. However, in either case, someone must pay attention to incentives and, in effect, try to manipulate them toward desired ends. This is not a maneuver that one level of government can perform very effectively on another--or on itself. This is an important role for the replication agency, as will be discussed later.

Complex as the incentive issue is, replications do have an edge on other types of innovation, specifically because they are not brand new. That is, by the time a program gets replicated, debate over its purposes, if not ended, is likely to be considerably diminished. Evaluation research will have established fairly clear output measures. Implementational research and systematic reflection on the experience of pilot programs should also provide fairly clear standards of practice and performance for program staff.

For these reasons, it seems intuitively probable that replications have a better chance of establishing something like a "bottom line" than programs whose purposes are more subject to different political or professional interpretations. Because the performance measures for replication emerge from practice, not from a regulator's best guess, they have greater face validity and it is easier--at least, intellectually--to insist on accountability for achieving them. And, finally, because something is known about the programming standards and professional practices associated with results, technical assistance can be authoritative and tied to the incentive of achieving results.

3. The Replicating Agent

A third basic success factor present in the five examples is an organization responsible for carrying out the replication: a replicating agent. Change is rarely achieved without a change agent. In each case, the replication agent had a high stake in the success of the replication venture: it had developed the model.

The difficulties in making the transition from research to planning and from planning to implementation are widely recognized and well-documented. These are difficult transitions to make within a bureaucracy. The difficulty is exacerbated when the transition involves several levels of government, as the implementation literature that grew up dissecting Great Society programs amply documented. When these transitions involve replication, an entrepreneurial activity, it is particularly unreasonable to expect that government agencies will be good replication agents.

Each of these difficulties is at least partially addressed by vesting responsibility for replication in an intermediary organization whose business it is to make the transfer of knowledge succeed; which combines the skills and capacities of program development, research, training and technical assistance; and which can work across jurisdictions, local agencies and levels of government, attempting to harmonize a variety of incentive systems, without any turf entanglements or bureaucratic axe to grind. Of course, the technical skills and judgment of the agency's staff become important, but here too the use of intermediaries who have gone through the learning curve needed to develop these skills and select the right strategy for replication recommends itself.

In short, the nature of the replication enterprise suggests the importance of a replicating intermediary, whose presence answers many of the questions the implementation literature raises, at least implicitly, about the systematic application of social science to local program improvement. I do not think it an accident or a fluke of this study's sample selection that each replication reviewed employs an intermediary that at least generally exhibits all the attributes listed earlier. Further, just as Burger King needs its corporate headquarters, replications of social programs may need a replicating agent on a continuing basis; e.g., when the federal labs got out of the EBCE, EBCE died.

It is notable that the replicating agents in the examples reviewed here were all the developers as well as the reproducers of their programs. It is thus tempting to deduce that the best replicators will be the program's originators. There is some face plausibility to this conclusion: the author of a program may have a special sense of stewardship and commitment to its continuing success, as well as a well-honed sense of its essential features. But it is not clear that the few examples in this paper illustrate an invariable law. JAG and CCP were both based in part on existing programs or materials. P/PV as an organization was established initially under the YEDPA legislation to identify and replicate an "exemplary" community conservation and improvement program, and its VICI model was based in considerable degree on a Portland, Oregon program.

The distinguishing characteristic of the replicating agents in these examples may not be that they authored the programs. Rather, it may be that they are among a relatively small set of organizations designed and funded with the capacity (and sometimes the explicit mission) to execute all the tasks required in replication: i.e., research, planning, implementation, and technical assistance across levels of government and jurisdictions. Further, to grow and thrive as organizations, they have incentives to replicate and, since they live and die by the grant, considerable incentive to replicate well. (Note that in this latter sense, the nonprofit intermediaries differ from the regional education laboratories, whose interest in replication was closely tied to the mandate of their federal funds.)

In the private sector, analogously, it was McDonald's, not Burger King (or Wendy's or Roy Rogers), that invented the formula for the fast-food franchise. Indeed, the private sector is replete with examples of copy-cat firms that seek to apply someone else's formula for success. (Consider the mergers of insurance companies and brokerage houses into diversified financial service institutions, the IBM clone companies, the spread of law firms and accounting firms into management consulting, or the proliferation of sport coupes and Jeep-like vehicles among auto manufacturers.)

It is certain that there are many more examples of effective program models scattered across the country (and potential "replicatees") than there are agencies with the capacity and incentives to replicate (ergo, the widespread if generally undocumented phenomenon of locally initiated replications discussed earlier). But local government program operators have no mandate, time, funding or incentive to replicate elsewhere. Accordingly, it seems plausible that investment in organizations funded to identify and, as needed, research good programs--and follow through with further development and replication--could produce substantially more widespread use of replication as a strategy for the improvement of local practice.

THE CHOICE OF REPLICATION STRATEGY

Replication strategies, I have suggested, can be arrayed along a continuum from "high-fidelity, highly specified" approaches that are "product-oriented," to more adaptive, locally specified "process-oriented" approaches. Note that even high-fidelity approaches will require some adaptation--cloning is not possible --and a process for replication. "Process-oriented" replications will require fidelity to a model at some level of generality (e.g., key principles), or there will be no essential integrity. Further, process models ultimately require specification of program activities, but the responsibility for that specification lies primarily with the local agency or local professionals whom the process attempts to engage--and only secondarily with the replicators. There are, of course, a number of points along this continuum of replication strategies, as suggested by our examples. But to clarify strategic choices, it is helpful to examine the extremes.

The more "product-oriented" replications--CCP, JAG and STEP--have several commonalities. Each is a fairly highly specified program, replicated in a process that emphasizes fidelity to the model. Each, also, is a model with fairly easily quantified objectives, and research that demonstrates the model's success in meeting those objectives. None is so much a design for changing or confronting existing institutions as it is a discrete product that aims at a new or uncrowded market. Further, none flies in the face of conventional wisdom. Indeed, there is fairly broad consensus, in the research and among professionals, about the appropriateness of the types of services CCP, JAG and STEP offer for the target populations they serve.

The more "process-oriented" replications we have examined, EBCE and CES, were designed to be more highly adaptive. The goals of these efforts are somewhat broader and more diffuse than those of the first three models. Neither is designed for a specific target population; rather, each is a vision of education for all students. While EBCE was undergirded by experimental research, the types of indicators on which the program was judged successful tended to be soft: e.g., students' self-reported interest in school and confidence in their ability to achieve career goals. CES depends primarily on systematic professional inquiry and speculation, for its underpinnings rather than quantitative research. The institutional environments both inhabit are considerably different from those of CCP, STEP and JAG. The education world is less results-oriented than the job training world, or the private sector realm in which JAG partially operates.

Both EBCE and CES are squarely located in the bureaucratically dense, professionally entrenched world of school systems. They are attempts to change both the structure of those institutions and the day-to-day behavior of professionals. EBCE sidestepped somewhat the issue of institutional change by setting itself up as an alternative program, but the alternative nevertheless represented significant changes in institutional practice. CES challenges the institution head-on. The ambitiousness of these efforts--part and parcel of the diffuseness and hard-to-quantify nature of their goals--makes hard, quantifiable measures inherently more difficult to achieve. Further, their ambitious change agenda, the softness of their research and the dense, professionalized settings they seek to change require that they recognize the need for "mutual adaptation." Modifying the model may be the price of modifying professional practice, through the process of mutual adaptation, which the Rand educational change study (and much other research on the implementation of innovations in the late 1960s and early 1970s) found to be characteristic of successful innovations (Berman and McLaughlin, 1974).

In fact, while CCP, JAG and STEP make a pretty strong case for the viability of "product-oriented" replications that fill a market niche, neither EBCE nor CES clearly demonstrates the viability of replication as a strategy for institutional change. EBCE succeeded in replicating an alternative program, but in so doing it finessed the issue of broader change. CES has a strong change agenda--but it is too early to say whether it will prove itself in more than a handful of situations.

In sum, these examples and intuition suggest that both the nature of the model and the environment within which it is to be replicated will influence the choice of replication strategy. High-fidelity, highly specified, "product-oriented" models may work in providing a new service or filling a market niche, in cases where there is little bureaucratic or professional resistance. Adaptive, loosely specified, "process-oriented" strategies are required when the goal is institutional change in bureaucratically and professionally dense environments. Further, the examples we have reviewed are grounded in a general vision of what education should be. Their visions require significant changes in both the organization of school and the activities of teachers, but they allow substantial room for schools and teachers to shape the specifics of the approach.

While the examples we have used tend fairly strongly to cluster the types toward one end of the continuum or the other, mixed-mode combinations of strategies are clearly possible. For example, one would expect that even a limited program model, with no pretensions to institutional change, would need to leave room for adaptation if it requires changes in bureaucratic or professional habit. Similarly, P/PV's STEP program uses a high-fidelity approach in the relatively simple environment of summer jobs, and a highly adaptive approach in implementing the school year support component.

TECHNICAL TASKS IN REPLICATION

Once it has developed credible evidence; identified a market, the incentives, and the strategy; and either bought or sold participating sites (often, some combination of the two), the replicating agency gets down to work. Here, the primary tasks are fairly generic to much good program implementation. They include site selection, training and technical assistance, and ongoing quality control or monitoring of results.

By and large, there is nothing unique about good practice in performing these tasks in replication efforts, but there are a few issues that, based on what replicators told me, deserve special attention.

Site Selection

Competition, self-selection or persuasion (coercion) are the possible methods of site selection. Persuasion is the least likely to succeed, since it is hard enough to replicate without also having to lead a reluctant horse to water.

Both competition and self-selection permit careful screening of sites by the replicating agency, just as Burger King carefully screens prospective franchise owners and locations. Such screening is particularly important to assure that the local boundary conditions (e.g., funding arrangements) and informed commitments are in place (e.g., "champions" in influential places, teacher support).

Replicators differ, however, concerning the degree to which one can predict success from initial screening, and thus the amount of effort that should go into it. Some staff at U.S. BASICS and P/PV, for instance, question whether they need to put as much effort as they do into this stage--perhaps because they also provide more intensive technical assistance and/or have stronger quality control over program planning.[24] By contrast, Hagans reported that in EBCE almost no one was turned down, reflecting NWREL's endorsement of high adaptability as well as the political problems it causes a regional lab to say no to an LEA. (This may well explain why Farrar et al. likened EBCE to a garden party.)

As a rule, however, careful selection seems critical to achieve both local administrative pressure for implementation, and good program "fit" with the local situation. Both administrative pressure and fit are crucial to replication, and both emerge as critical success factors in the change literature (McLaughlin, 1987; Huberman and Miles, 1984).

Training and Technical Assistance

CES, EBCE, CCP and STEP have all found it quite beneficial to use staff from other successful sites in providing planning and technical assistance--for much the same reason JAG found it effective to have governors sell the program to other governors: high credibility. The incorporation of peers seems particularly suited to programs operating in professionalized, bureaucratic environments, where there is a pronounced tendency to be skeptical of people who "haven't done it" or "don't know what it's like." This approach is less appropriate to JAG, whose affiliates hire their own school-based staff, and CCP, whose franchise holders possess a less stable and professional work force.

Both CCP and STEP found strong appetites for detailed curricula and training manuals. P/PV's Jerry Kolker remembers that in the course of the demonstration phase, "teachers got furious at us because we expected them to develop the curriculum." In part, this no doubt reflects the fact that most teachers had little time, resources or experience for developing the kind of curriculum P/PV required for a summer program. But it may also reflect the lack of "habit" for thinking about doing things differently, which Sizer noted. Even in the process-oriented world of CES, Sizer reports, teachers flock to other teachers at workshops for concrete examples of how to do it, despite the fact that CES strongly encourages people to "do it your way." (Starting something new, most of us seek precise instructions. First steps are usually highly imitative--another fact consistent with the literature on change.)

Whatever its precise form, there can be little doubt that training and technical assistance--habitually underfunded public activities--are critical to the replication process, and to change itself. In their impressive study of innovation in schools--many of them small-scale replications funded by the U.S. Department of Education's National Diffusion Network--Huberman and Miles (1984) stress that complex innovations "live or die" based on the quantity and quality of training and technical assistance provided once the change process was begun.

Quality Control

Ongoing quality control can be problematic, especially if there is no strong franchise agreement. If the best test of maintaining the essential integrity of a model is the ability to replicate results, ongoing evaluation is important. But strong and comprehensive research designs are expensive. They diminish in feasibility as the instances of replication multiply. (This problem is not unique to replication strategies, of course, but it is more central to their raison d'etre.)

A continued role for the replicating agency in collecting and analyzing performance data is the most obvious response to this problem, but Rex Hagans had a useful suggestion for an alternative approach: stress local responsibility for evaluation during site development, with provision of needed technical assistance, to assure site "ownership" of the evaluation process, as well as of the program. Then, let the replicating agency or a third-party evaluator make periodic audits of the quality of local evaluations. This approach has appeal both as a means of assessment and as a way to stimulate continuing evolution and improvement of the program after the replicator is gone. Further, in school settings particularly, it gains support from a considerable body of literature suggesting that research is most likely to be used by professionals when they have a hand in generating it (Lieberman, 1986; Evertson, 1986; Cousins and Leithwood, 1986).

There is another dimension to the issue of quality control. If replication is an entrepreneurial activity, and if it is best done by a replicating agent that combines research and implementational skills, the potential of a conflict arises. That is, will the replicating agent's entrepreneurial aims lead it to substitute hype for carefully assessed hypotheses in researching the model or the continued effectiveness of the replications? There is a clear and present danger here, but not an insurmountable flaw. First, so long as the replicating agency requires some external support to do its work (and all the examples cited suggest that the need for such support continues for a long time), funders will be in a position to review the quality of the agency's research--or even hire independent evaluators--to assure accountability. Second, the incentives for distorting research are no greater for a replicating agency than for any operating entity. Indeed, they may be less, for I would venture that in the long run, the credibility of the replicating agency, with both funders and local programs, will be directly tied to its success.

PROMOTING REPLICATION: FEDERAL, STATE AND FOUNDATION OPTIONS

This paper has argued that planned replication in most instances is best carried out by entrepreneurial organizations that can operate across various levels of government and agency jurisdictions over long time periods. Neither the legislative nor the administrative branches of federal or state government are well positioned to be the replicating agency in most instances. However, these superordinate governmental levels, as well as major foundations, have potentially important roles to play in stimulating and supporting replication. They are the most likely supporters of research designed to produce credible knowledge about "what works," and they are likewise best positioned to capitalize replicating agencies. They can help create local receptivity to replication by partially defraying costs of local program operation. Less directly, legislation grounded in well-conceived performance objectives also appears, on the face of it, likely to create incentives which increase the likelihood of local programs seeking "best practices."

The examples we have reviewed provide illustrations of these roles. In EBCE, the federal government underwrote knowledge development, supported the replicating agencies and seeded the market with funding for local programs. The Department of Education's National Diffusion Network remains an intriguing approach to tapping decentralized creativity in developing credible knowledge, by validating research initiated by local scholars and practitioners, although it has lacked the funding to follow through with whole-scale replication. Both STEP and CCP were heavily capitalized by major foundations, and JTPA performance regulations and programming requirements helped define their market niches. JAG, developed with support of foundation and federal dollars, has benefited in some states from legislative appropriations that have speeded its growth.

With the exception of EBCE, none of these examples represents full-scale, start-to-finish replication strategies planned and carried out by a governmental agency over a period of years, and EBCE itself withered away when the regional labs--which were grant-driven, not market-driven, entities--lost funding for it. The paucity of examples of government-supported replication no doubt reflects a combination of factors, including limited technical knowledge about how to replicate and, more important, lack of a political constituency (local entities want the funds with as few strings attached as possible) and the incompatibility of replication's long timelines with the shorter horizons of electoral politics.

The technical knowledge exists, this paper suggests, but the political obstacles are harder to overcome. Nonetheless, an appealing case can be made for federal or state support of replication on grounds that, despite up-front research and development costs, it offers long-term possibilities for leveraging federal (or state) dollars both directly and indirectly. Directly, in that the market nature of replication, as we have defined it, argues that cost-sharing or matching requirements should be integral to support of replicating agencies and/or local program operators. Indirectly, in two senses: first, successful replications promise to increase the effectiveness of local expenditure of formula funds, and second, skillful replicating agencies should be able, as some of our cases have shown, to pool funds from a variety of sources at the local level, creating synergies that task forces on interagency cooperation often envision but less frequently achieve.

If the federal government, in particular, is to stimulate replications in education and employment training, the ideal approach might well be to borrow from other fields, such as the National Institutes of Health or the National Science Foundation, where semi-autonomous entities have been fairly successful in implementing long-term R & D programs with minimal political disruption. If this sounds like a reinvention of the National Institute for Education, it is, but possibly the climate is more hospitable now for a strong federal R & D presence, and an institute devoted to the improvement of local programming would eliminate one of the most common objections to the old NIE--that its research had no benefit to local practitioners. There would be good logic, flowing from the arguments of Workforce 2000 and similar studies which have found a receptive audience in both industry and government, to a combined National Institutes for Education and Work, which would work closely with both the Department of Labor and the Department of Education, and whose agenda would include not only improving basic education and job training for the disadvantaged, but also broader issues of workforce literacy and upgrading. Such an entity would be charged with developing and sustaining the replication agenda--from R & D through adaptation. It would oversee funding and performance of replicating agencies. It could also provide incentive funding to seed local adaptations, or it could work with Labor, Education and other departments to make such funding available through their budgets.

More modestly, without creating any new entity, renewed attention by federal (and state) governments to building replication capacity and routinely linking replication efforts to program evaluation and research, whether through existing entities like the regional laboratories or through competitive support of a wide range of nonprofit replicating agencies, is critical if replication is to recognize its potential as a strategy for improving implementation.

V. CONCLUSIONS: WHEN TO REPLICATE AND WHY

Planned replication is not the only strategy for improving local programming. Sensible legislation, intelligent regulations that stress outcomes and restrain the impulse to prescribe inputs, adequate funding, improved pre-service and in-service training, good management, restructured school governance, efforts to stimulate competition among service providers (e.g., choice plans) are among other strategies whose importance is self-evident.

Replication is not a panacea, and there are many qualifications governing its effectiveness, as this paper has been at pains to point out. Benefits from replication are likely to be modest and incremental, just as the benefits of research-validated models are almost always modest. But this is in reality a strength, not a criticism. Studies of research utilization establish that research is most likely to influence practice when it does not depart dramatically from what is widely believed and when, indeed, it tends to confirm much of what people already know (Cousins and Leithwood, 1986). In this, it resembles changes in scientific knowledge as well as the progress of social science (Lindblom and Cohen, 1979). Anything else is a revolution.

Planned replication need not compete with other strategies for improving implementation. Rather, it complements them. A sensible, adequately funded piece of legislation, such as Head Start, could yet benefit from planned replication. A highly decentralized effort to restructure schools may benefit from replications when it comes time to answer the question of replication's educational purpose. In such instances, the restructuring may be a "process-oriented" replication, or it may be that a group of teachers, once empowered, will seek a program model to address a particular aspect of school improvement.

This paper suggests that "product-oriented" replications are more proven, while more visionary "process-oriented" replications are more problematic. I suspect this is true: that replication has more potential in discrete program applications, occupying a market niche, than for broad institutional change. The reason lies not only in greater complexity, but also in the fact that the more a third party governs process, the more it tends to usurp functions of the local administrative hierarchy. However, this cautionary note should not rule out the possibility of successful process-oriented replications. Rather, it points out the great importance of support from the top--even for "bottom up" change strategies (e.g., CES requires local school boards to give schools great freedom from central dictates for at least three years).

More positively, market-oriented replications as I have defined them may represent at least a partial answer to some of the dilemmas of improving local practice in our decentralized, federalist system of government. By its very nature, it includes valid, practice-forged outcome measures that are necessary (if not completely sufficient) to inculcate a "results orientation" in local programming, and it offers a technology, through the actions of the replicating agent, not only for moving research into practice but for tying program design and staff training to results.

In the 1970s, some were led to question the ability to apply social science to improving social programming. In the wake of the Great Society, and the lesser aftermath of such efforts as the Head Start and Follow Through Planned Variations, a swelling literature on implementation raised substantial doubts about the ability of federal initiatives to influence local practice (e.g., Murphy, 1971; Pressman and Wildavsky, 1974; Berman and McLaughlin, 1974; Williams and Elmore, 1975). One strand of thought, epitomized by Wildavsky, emphasized the complexities, tensions and diversity of conflicting interests that confound rational implementation on its journey through public bureaucracies. Another strand, epitomized by McLaughlin in the Rand change studies, was derived from studies of organizational behavior in the corporate world. It emphasized not only the need to change whole institutions, but the need to engage the staff of institutions in the planning and implementation of change through a process of "mutual adaptation." This strand of thought is alive and well in the emphasis on what is variously called "school site management" or "restructuring" or "teacher empowerment," which is the cresting wave of educational reform today. Peter Drucker's dictum, "there can be no divorce of planning from doing knowledge work," is an apt motto for this vision of school reform, which some suggest is antithetical to the idea of replication.

Replication is, to be sure, a rational form of implementation, and schemes like the Head Start Planned Variations that imply a top-down and more-or-less universal application of knowledge are probably doomed to fail. But thinking of replication as an entrepreneurial, "market-oriented" process avoids the pitfalls of top-down, intergovernmental implementation schemes. Similarly, the "process oriented" replications represented by EBCE and CES appear to be highly congruent with research on the need for basic institutional change and "mutual adaptation" in complex bureaucracies, such as school systems.

REQUISITES FOR REPLICATION

It is fair, then, to ask how widely applicable planned replication is as a strategy, and what conditions are necessary to permit it to occur. The answer, like the process, has several parts. There are a limited number of program models that have developed credible evidence of effectiveness--whether that evidence is produced by strict experimental means or by other forms of systematic assessment. It is notable that the examples cited in this paper constitute about half the "proven" initiatives identified by Berlin and Sum (1988) in their review of education and job training approaches effective for disadvantaged youth and young adults. One may find the paucity of proven models discouraging--or be encouraged by the fact that such a high percentage have been replicated. In either case, however, planned replication, as I have defined it, will have a small future without continued efforts to develop and research effective models with a clearly understood essential integrity.

To make the point more sharply, the apparent lack of programs worth emulating may well say more about the paucity of investment in efforts to identify promising examples and to develop credible evidence of their effectiveness (and viability in multiple settings) than it does about the lack of effective programs per se. Casual empiricism--i.e., my own observations over the years --suggests that this is in fact the case. In my experience, at least, many a flower blooms and dies unnoticed in the ill-tended garden of public programs.

Assuming continued development of the knowledge base, replication makes sense when the following conditions are met:

1. There is credible evidence on the basis of which one can reasonably infer the program model is both effective in meeting its objectives and at least as effective as any known alternative on which funds would otherwise be spent.

2. Input costs of the model are not out of line with prevailing cost standards.

3. A replicating agent with the requisite technical skills to market and reproduce the model can be found.

4. There is an identifiable market niche, either for a particular program or service in an institutionally weak field, or for more basic organizational change. Here, the best proof is the willingness of local agencies to put their money on the table.

5. The choice of replication strategy is appropriate to the product and the institutional context in which the replication will occur.

6. There are reasonable provisions for continuing evaluation of effectiveness.

The potential sphere for planned replications, assuming credible evidence can be developed, strikes me as quite large. Abstractly, it includes the whole set of programs that exist or might exist in market niches, or for which practice is currently weak, in addition to possible initiatives for basic institutional change. More concretely, I would argue that strong needs for better models (or better researched models)--and a prima facie case for replication--exist in a wide range of areas. For example:

- Early childhood education, where abortive initial attempts to conduct planned variations have not vitiated the desirability of finding out what works (perhaps different models for different purposes) and disseminating it;
- Special education, where the political pressures for mainstreaming exceed the knowledge of how to do it effectively;
- Vocational education, where about all we know is that the traditional model does not work very well (and where some interesting sounding alternatives flourish largely in isolation, with little systematic evidence to validate their effectiveness);[25]
- Vocational education for special education students--required everywhere by law and done almost nowhere with effectiveness;
- Models for blending remedial education and job training in an integrated curriculum--performed with some success in the military (Sticht et al., 1987) but little attempted in the public sector; and
- Adult literacy programs, where models capable of holding more than a handful of participants long enough to yield benefits appear to be scarce or nonexistent.

One could go on. Suffice it to say, the conceptual possibilities are immense, and the technology, so primitive only 20 years ago, is much better developed today.

One further condition must be met, however, to capitalize on the possibilities--and that is, in a more traditional sense, the problem of capital. All replications require front-end development costs, and the replications we have examined required ongoing grants to provide training, technical assistance and other services of the replicating agency. JAG and CCP have covered some of these costs through fees charged to their franchisees. Some of P/PV's STEP replication costs are offset by grants from foundations and federal and state agencies. But all need some combination of venture capital and working capital. In the public and nonprofit sectors, government and foundation grants are, of course, the only source of such capital. From development to current status, all of the replications we have reviewed are at least five years old, and all continue to require subsidy. Funders interested in pursuing replication must be prepared to invest for the long haul.

A reasonable person may wonder: Is it really worth it? So what if a replication reaches 200 sites or 30,000 people before it runs out of steam? So what if the program is actually institutionalized in those sites? Is that enough to really make a difference? Does it have to go on longer? Assuming all formal replications must end, will the informal processes of knowledge diffusion and technology transfer, however hard to trace, actually take hold? Are the returns great enough to justify long-term investment in the replication process?

The honest answer, it seems to me, is that no one really knows. But perhaps the burden of proof lies elsewhere: if not by replication, then how?

NOTES

1. A computer search of ERIC, PAIS, Social Science Abstracts, the Educational Index, Index, INFOTRACK, the Government Publications Index using the terms "adaptation", "replication", "demonstration", "diffusion", "dissemination", "program development", "program implementation", "program validation", "technology transfer" and "research utilization" turned up only a handful of relevant entries.

2. P. Michael Timpane, president of Teachers College, Columbia University, is the source of this quip, commenting on the reaction of teachers in New York City he encountered while exploring the possibility of an effort to identify and replicate successful programs there during 1988.

3. Gary Walker, executive vice president of Public/Private Ventures, suggested this definition.

4. Initially, Follow Through was planned as a major piece of funding to support programs which would supplement Head Start in the early grades. When it became apparent that Congress would not fund a major new initiative, its backers settled for a demonstration program, the Planned Variations experiment.

5. Telephone interview by the author with P. Michael Timpane, who ran NIE during this transition of emphasis.

6. For example, Schweinhart and Weikart criticize McKey et al. for including all available studies in their meta-analysis, instead of limiting the sample to the best studies. But McKey reported having rejected this alternative after a preliminary analysis showed that limiting the sample in this fashion would have produced weaker Head Start benefits.

7. Telephone interview by the author with Theodore B. Sizer.

8. Telephone interview by the author with Marion Pines.

9. Author's interview with Joseph Seiler, vice president, U.S. BASICS.

10. Telephone interview by the author with Ben Lattimore, OIC.

11. Author's interview with Jerome Kolker, P/PV's vice president.

12. Telephone interview by the author with Susan Bloom of the STEP replication staff.

13. Most research has failed to show a connection between knowledge of birth control and sexual behavior.

14. Telephone interview by the author with Ken Smith.

15. Process described by Ken Smith in telephone interview.

16. Telephone interview by the author with Andrea Baker, Northwest Regional Laboratory.

17. Telephone interview by the author with Rex Hagans.

18. Ibid.

19. Telephone interview with Sizer.

20. Ibid.

21. Howard Rosen, who directed much of the U.S. Department of Labor's major demonstration research in the 1970s, believes that the research did influence congressional policy, albeit in general ways. (Interview with author.)

22. Telephone interview by the author with Mitchell Sviridoff.

23. Gary Walker, who played key roles throughout the Supported Work demonstration as executive vice president of MDRC, is the source of this information.

24. Interviews with Joseph Seiler and Jerome Kolker of U.S. BASICS and P/PV, respectively.

25. Large-scale studies of vocational education have generally failed to show lasting labor market benefits for students, although these studies, such as Grasso and Shea's work for the Carnegie Council on Higher Education in the 1970s and the NIE study in the late 1970s, have been limited in their ability to control for quality of program. On the other hand, some studies of discrete programs, such as P/PV's analysis of the Philadelphia High School Academies program, have found positive impacts on graduation rates and employment. See P/PV, Allies in Education, 1987.

REFERENCES

Berman, P., and M.W. McLaughlin. Federal Programs Supporting Educational Change, Vol 1: A Model for Change. Santa Monica, Cal.: Rand Corporation, 1974.

Berrueta-Clement, J.R., et al. Changed Lives: The Effects of the Perry Pre-School Program on Youths through Age 19. Ypsilanti, Mich.: High Scope Educational Research Foundation, 1984.

Berlin, G., and A. Sum. Toward a More Perfect Union: Basic Skills, Poor Families and Our Economic Future. New York: Ford Foundation, 1988

Betsey, Charles; Robinson Hollister and Mary Papageorgiou. Youth Employment and Training Programs: The YEDPA Years. Washington, D.C.: National Academy Press, 1985.

Brandt, R. "On Changing Secondary Schools: A Conversation with Ted Sizer," Educational Leadership, February 1988.

Clune, W.H., et al. The Implementation and Effects of High School Graduation Requirements. New Brunswick, N.J.: Center for Policy Research in Education, February 1989.

Coalition of Essential Schools. Prospectus. Providence, R.I.: Education Department, Brown University, 1988.

Comer, J. Jr. School Power: Implications of an Intervention Project. New York: The Free Press, 1980.

Committee on Economic Development. Children in Need: Investment Strategies for the Educationally Disadvantaged. New York and Washington, D.C.: CED, 1987.

Cousins, J.R., and K.A. Leithwood. "Current Empirical Research on Evaluation Utilization," Review of Educational Research, res util, 1986.

Datta, Lois-Ellen. "A Report on the Evaluation Studies of Project Headstart," International Journal of Early Childhood, Vol. 3, No. 2, pp:58-69, 1971.

Elmore, Richard F. "Design of the Follow-Through Experiment," in Planned Variation in Education: Should We Give Up or Try Harder?" A. Rivlin and P.M. Timpane, eds. Washington, D.C.: The Brookings Institution, 1975.

____________. Early Experience in Restructuring Schools: Voices From the Field. Washington, D.C.: National Governor's Association Center for Policy Research, 1989.

Evertson, C. "Creating Conditions for Learning: From Research into Practice," Theory Into Practice, Vol. XXVI, No. 1. January 1986.

Farrar, E.; J.E. De Sanctis and D.K. Cohen. Views from Below: Implementation Research in Education. Cambridge, Mass.: The Huron Institute, 1979.

First Pennsylvania Bank. "Business and Education: No Longer Separate Dimensions." Regional Reports. First Quarter, 1989.

Grubb, N., and M. Lazerson. "Rally Round the Workplace: Continuities and Fallacies in Career Education," Harvard Educational Review, November. 1975.

Huberman, A.M., and M.B. Miles. Innovation Up Close: How School Improvement Works. New York: Plenum Press, 1984.

Lieberman, A. "Collaborative Research by Teachers and Principals," The Education Digest, May 1986.

Lindblom, C.E., and K.K. Cohen. Usable Knowledge: Social Science and Social Problem Solving. New Haven: Yale University Press, 1979.

Mallar, C., et al. Evaluation of the Economic Impact of the Job Corps Program, Third Follow-Up Report. Princeton, N.J.: Mathematics Policy Research, 1982.

Manpower Demonstration Research Corp. Summary and Findings of the National Supported Work Demonstration. Cambridge, Mass.: Ballinger, 1980.

Marland, S. P. Jr. Career Education. Washington, D.C.: Office of Education (DHEW), September 1971.

McKey, R.H., et al. The Impact of Head Start on Children, Families and Communities. Washington, D.C.: Government Printing Office, 1985.

McLaughlin, M.W. "Implementation as Mutual Adaptation: Change in Classroom Organization," Teachers College Record, February 1976.

___________. "Learning from Experience: Lessons from Policy Implementation," Educational Evaluation and Policy Analysis, Summer 1987.

Mosteller, F. "Comment" in Planned Variation in Education: Should We Give Up or Try Harder?, A. Rivlin and P.M. Timpane, eds. Washington, D.C.: The Brookings Institution, 1975.

Murphy, J.T. "Title I of ESEA: The Politics of Implementing Federal Education Reform," Harvard Educational Review, February 1971.

National Commission on Excellence in Education. A Nation at Risk. Washington, D.C.: U.S. Department of Education, 1983.

Owens, T.R. "Experience Based Career Education: Summary and Implications of Research and Evaluation Findings," Child and Youth Services Journal, Vol. 4, No. 3/4, 1982.

Powell, A.; E. Farrar and D.K. Cohen. The Shopping Mall High School: Winners and Losers in the Educational Market Place. Boston: Houghton Mifflin, 1985.

Pressman, J., and A. Wildavsky. Implementation. Berkeley, Cal.: University of California Press, 1974.

Public/Private Ventures. Summer Training and Education Program: A Preliminary Report on the Summer 1985 Experience. Philadelphia: Public/Private Ventures, 1985.

Public/Private Ventures. The Longer Term Impacts of Pre-employment Services on the Employment and Earnings of Disadvantaged Youth. Philadelphia: Public/Private Ventures, May 1983.

Rivlin, A. Systematic Thinking for Social Action. Washington, D.C.: The Brookings Institution, 1975.

Rivlin, A. and P.M. Timpane, eds. Planned Variation in Education: Should We Give Up or Try Harder? Washington, D.C.: The Brookings Institution, 1975.

______________. "Planned Variation in Education: An Assessment" in Planned Variation in Education: Should We Give Up Or Try Harder?, A. Rivlin and P.M. Timpane, eds. Washington, D.C.: The Brookings Institution, 1975(b).

Schorr, L.B. Within Our Reach: Breaking the Cycle of Disadvantage. New York: Anchor Books, 1988.

Schweinhart, L.J., and D.P. Wiekart. "What Do We Know So Far? A Review of the Head Start Synthesis Project," Young Children, January 1986.

Silberman, C. Crisis in the Classroom. New York: Random House, 1971.

Sipe, Cynthia; Jean Grossman and Julita Milliner. Summer Training and Education Program: Report on the 1987 Experience. Philadelphia: Public/Private Ventures, 1988.

Sizer, T.R. Horace's Compromise: The Dilemma of the American High School. Boston: Houghton Mifflin, 1984.

Smith, Thomas J. The Private Sector Initiative Program: Demonstration and Assessment of CETA Title VII Implementation: Final Report. Philadelphia: Public/Private Ventures, 1982.

Snyder, Phyllis, and Bernard McMullan. "A Profile of Philadelphia High School Acadamies," in Allies in Education: Volume 2. Philadelphia: Public/Private Ventures, 1987.

Spotts, R.L., and B.B. Hampson. National Directory of Experience-Based Career Education Programs. San Francisco: Far West Laboratory for Educational Research and Development, 1979.

Sticht, T.G. et al., Cast-Off Youth: Policy and Training Methods from the Military Experience. New York: Praeger, 1987.

Sum, A. The Longer-Term Impacts of Participation in Jobs of America's Graduates Programs: Findings of the 20-Month Follow-Up of the Class of 1984. Washington, D.C.: Jobs for America's Graduates, 1987.

Sum, Andrew, and Robert Taggart with Gordon Berlin. "Cutting Through," paper prepared for The Ford Foundation Project on Social Welfare and the American Future, 1986.

Taggart, R.A. Fisherman's Guide: An Assessment of Remediation and Training Strategies. Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research, 1981.

________________. The Comprehensive Competencies Program: Statistics, Second Quarter, 1987. Washington, D.C.: Remediation and Training Institute, August 1987.

Timar, T. "The Politics of School Restructuring," Phi Delta Kappan, December, 1989.

Walberg, Herbert J., et al. "Reconstructing the Nation's Worst Schools." Phi Deltas Kappan, Vol. 70, No. 10, pp: 802-05, June 1989.

Walker, G., et al. An Independent Sector Assessment of the Job Training Partnership Act. New York: Grinker, Walker & Associates, 1985.

Watkins, R., and R. Corder. Student Outcomes and Participant Opinions in Experience Based Career Education Schools, Vol. VI. Berkeley, Cal.: Educational Testing Service, 1977.

Weick, K.K. "Education Organizations as Loosely Coupled Systems," Administrative Sciences Quarterly, March 1976.

Williams, W., and R.F. Elmore. (eds.) Social Program Implementation, New York: Academic Press, 1976.

Williams, W. "Implementation Problems in Federally Funded Programs," in Williams and Elmore, eds. Social Program Implementation, New York: Academic Press, (1976).

U.S. Department of Education. What Works: Research About Teaching and Learning. Washington, D.C.: Government Printing Office, 1987.

________________. What Works: Schools that Work: Educating Disadvantaged Children. Washington, D.C.: Government Printing Office, 1987.